GRADUATE SCHOOL APPROVAL RECORD
NORTHEASTERN UNIVERSITY

Dissertation Title: The Development and Application of a MALDI-TOF MS Method for Quantitative Glycomic Analysis

Author: Samnang Tep

Department: Chemistry and Chemical Biology

Approved for Dissertation Requirements of the Doctor of Philosophy Degree

Dissertation Committee

Prof. William S. Hancock, Advisor 21 Nov 2011

Prof. Paul Vouros

Prof. Penny J. Beuning 18 Nov 2011

Zoran Sosic 18 Nov 2011

Head of Department

Prof. Graham Jones 11/12/11

Graduate School Notified of Acceptance

Amy Killeen 11/28/11
THE DEVELOPMENT AND APPLICATION OF A MALDI-TOF MS METHOD FOR
QUANTITATIVE GLYCOMIC ANALYSIS

A dissertation presented

by

Samnang Tep

to

The Department of Chemistry and Chemical Biology

In partial fulfillment of the requirements for the degree of

Doctor Philosophy

in the field of

Chemistry

Northeastern University
Boston, Massachusetts
November 2011
THE DEVELOPMENT AND APPLICATION OF A MALDI-TOF MS METHOD FOR
QUANTITATIVE GLYCOMIC ANALYSIS

By

Samnang Tep

ABSTRACT OF DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Chemistry and Chemical Biology in the

Graduate School of Science of

Northeastern University, November 2011
ABSTRACT

Glycosylation has been determined to be involved in essentially all aspects of life, ranging from development to ageing to health and to disease. In the last 25 years with the advent of recombinant biotherapeutics, it has even become pivotal in drug discovery and development. On the heels of genomics and the completion of the human genome, came the emergence of other spin-off or daughter “omics” fields. Included in this category is glycomics; the study of the “entire set of glycans in an organism”. With glycomic studies, came the observation that glycosylation changes are associated with various diseases and disease states. In response to these observations, methods were developed to characterize and quantitate these changes, which have led to advances in biomarker discovery. Meanwhile in the biopharmaceutical industry, numerous efforts were being made to control, characterize, monitor and quantitate product glycosylation. As often is the case, these efforts were seemingly separated by an imaginary line between academia and industry. For the most part, this is understandable as the needs of each “side” are inherently different. Academic pursuits are often lengthy and labor intensive resulting in highly detailed information, while industrial applications are often more focused on robust, rapid analysis that provides very specific types of information. The goal of this work was to erase this line by developing a method that is universally applicable to both academia and industry.

An innovative method utilizing well-established chemistries, commercially-available reagents and commonly-used instruments has been developed and applied to three different types of samples: (1) a purified biotherapeutic  (2) CHO cells (3) human plasma. The method employs reductive amination of PNGase F-released glycans with carbon-isotopes of anthranilic
acid and quantitative analysis by MALDI-TOF MS. The method displays high reproducibility (CV generally less than 10%) and data accuracy (generally 90% or greater) with a linear dynamic range over two orders of magnitude and sub picomolar sensitivity.

The method was applied to the analysis of a purified biotherapeutic over the course of a bioreactor campaign; six time points were sampled. The change in glycosylation over time (at each time point relative to the first time point) was quantitated, and it was observed that sialylation showed a marked decrease at the final time point. As sialylation often plays a role in the pharmacokinetic properties of biotherapeutics, this type of information may be of importance and could be used to modify cell culture parameters.

The method was also applied to the glycomic analysis of biotherapeutic-expressing CHO cells over the course of a bioreactor campaign; two different cell lines expressing two different monoclonal antibodies were sampled at four time points. The change in glycosylation over time was quantitated. Numerous glycans were observed; some increased over time, while others decreased. The impact of this information is unknown at this point and will require further investigation. The method was used in this application to demonstrate the feasibility of its use for quantitative glycomic investigations of CHO cells. In future experiments, it will be used to investigate whether CHO cell glycosylation has any impact on cell health and/or product quality.

The method was then applied to the glycomic analysis of normal vs. diseased human plasma. The method was combined with other preparative methods to establish a platform methodology for the purification and quantitative glycomic analysis of human plasma. Differences in glycosylation between normal and diseased plasma were observed and quantitated. Diseased plasma was observed be elevated in sialylation, by as much as 500% for a
particular glycan. The observation of increased sialylation was expected as other cancers have been observed to show increases in sialylation. The method was utilized in this setting to demonstrate proof of concept for its use in biomarker discovery.
ACKNOWLEDGEMENTS

The completion of this thesis is an accomplishment I could not have realized were it not for the support of my family, colleagues and friends. I would like to specifically recognize and thank the following people.

To my family, especially my mother and father who have sacrificed so much of themselves to ensure that my four siblings and I were appropriately provided for and had all the tools we needed to succeed in life. Thank you for risking your lives to flee your war-torn home country to seek a better life for your family in America. Thank you for working multiple jobs to provide a home for your family and thereby instilling in us all a work ethic, an honesty and an appreciation for life that is unparalleled. Thank you for supporting me in all of my endeavors, academic and otherwise. Thank you to all of my siblings (Rattana, Rattanak, Bophachan and Piseth) for supporting me in the ways that only siblings can. You all have helped to shape me into the person I am today. I love you all.

To my wife and best friend, Natalie. Thank you for all the support you have given me and the faith you have had in me throughout the years. Without you I would not have had the strength to accomplish this goal. Your warmth, guidance and encouragement has fueled me through the many challenges I have faced. The lessons you have taught me in patience and humility will undoubtedly help me to overcome the challenges I am sure to face in the future. I will always love you.

To my advisor, Dr. William S. Hancock. Thank you for your guidance and expertise. Thank you for giving me the opportunity to work on these interesting and exciting projects, and thank
you for believing in my abilities to accomplish this. The time and energy that you have invested in me has been an invaluable part of my growth as a scientist. Your enthusiasm and zeal are inspirational. For all of this, I will be forever grateful.

To my mentors/supervisors at Biogen Idec, Dr. Zoran Sosic and Joseph Siemiatkoski. Zoran, thank you for your continued support and guidance. Thank you for understanding the difficulties of completing a PhD program while working full time in a high-paced environment. Thank you for helping me to stay calm and focused. Joe, were it not for you I would have never entered the crazy field of carbohydrates and glycosylation. Thank you for setting the foundation for the scientist I have become. Thank you for believing in me and sharing with me your knowledge.

To my employer, Biogen Idec, and my management: Dr. Andrew Weiskopf, Dr. Yelena Lyubarskya, Dr. Helena Madden and Dr. Rohin Mhatre. Thank you so much for all of the support and encouragement you have provided me through the years.

To my lab mates at Northeastern and colleagues at Biogen Idec. Thank you all so much for your help, support and most importantly your friendship. A special thanks to Dr. Majlinda Kullolli, Dr. Agnes Rafalko, Dr. Marina Hincapie from Northeastern for all of the discussions and assistance. Also a special thanks to Dr. Damian Houde and Dr. Tyler Carlage from Biogen Idec for all your help with various projects throughout the years.

To my committee members, Dr. Paul Vouros, Dr. Penny J. Beuning and Dr. Zoran Sosic. Thank you for the time you have invested in me and the support you have given me. It will be appreciated forever.
Finally, thank you to my very close friends: Adam Noel, Beth McCooey, Thomas Snelling and Tyler Carlage. Thank you for always being there for me. Your support, encouragement and friendship are priceless to me. I love you all.
TABLE OF CONTENTS

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

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

TABLE OF CONTENTS .................................................................................................................9

LIST OF FIGURES .......................................................................................................................15

LIST OF TABLES .........................................................................................................................20

LIST OF ABBREVIATIONS ........................................................................................................21

CHAPTER 1

INTRODUCTION AND BACKGROUND: GLYCOSYLATION, ITS IMPORTANCE IN BIOLOGY AND BIOTECHNOLOGY, THE EMERGENCE AND EVOLUTION OF GLYCOMICS ..........................................................................................................................26

1.1 Glycosylation and Glycosynthesis ......................................................................................26

1.2 Glycosylation and Disease .................................................................................................31

1.3 Glycomics ............................................................................................................................37

1.4 Glycosylation and Biotechnology .......................................................................................65

1.5 Mass Spectrometry (Ionization) ..........................................................................................70
CHAPTER 2

DEVELOPMENT OF A MALDI-TOF MS METHOD FOR THE QUANTITATIVE ANALYSIS OF GLYCOSYLATION AND GLYCOMIC CHANGES

2.1 Introduction

2.2 Experimental

2.2.1 Reagents

2.2.2 Derivatization of Free Glycans with $^{12}$C$_{7}$ Anthranilic Acid and $^{13}$C$_{7}$ Anthranilic Acid

2.2.3 Deglycosylation of Glycoproteins

2.2.4 HPLC Analysis with Fluorescence Detection of Derivatized Glycans

2.2.5 Quantitative MALDI-TOF MS of Derivatized Glycans

2.3 Results

2.3.1 Premise and Methodology
2.3.2 Quantitative MALDI-TOF MS Analysis of Derivatized Glycan Standards

2.3.3 Quantitative MALDI-TOF MS Analysis of Derivatized Glycans from a Model IgG

2.3.4 Quantitative MALDI-TOF MS Analysis of Derivatized Glycans from a Model Fc-Fusion Protein

2.4 Discussion

2.5 References

CHAPTER 3

A MALDI-TOF MS METHOD FOR THE SIMULTANEOUS AND QUANTITATIVE ANALYSIS OF NEUTRAL AND SIALYLATED GLYCANS OF CHO-EXPRESSED GLYCOPROTEINS

3.1 Introduction

3.2 Experimental

3.2.1 Reagents

3.2.2 Deglycosylation of the Fc-Fusion Protein

3.2.3 Derivatization of Free Glycans with $^{12}$C$_7$ Anthranilic Acid and $^{13}$C$_7$ Anthranilic Acid

3.2.4 HPLC Analysis with Fluorescence Detection of Derivatized Glycans
3.2.5 Quantitative MALDI-TOF MS of Derivatized Glycans ........................................139

3.3 Results ...............................................................................................................................139

3.3.1 Quantitative MALDI-TOF MS Analysis of Derivatized Glycans from an Fc-Fusion Protein Over the Course of a Bioreactor Campaign ........................................................139

3.4 Discussion .........................................................................................................................144

3.5 References .........................................................................................................................145

CHAPTER 4

A MALDI-TOF MS METHOD FOR THE SIMULTANEOUS AND QUANTITATIVE ANALYSIS OF NEUTRAL AND SIALYLATED GLYCANS OF CHO-EXPRESSED GLYCOPROTEINS .......................................................................................................................150

4.1 Introduction .......................................................................................................................151

4.2 Experimental .....................................................................................................................152

4.2.1 Reagents .......................................................................................................................152

4.2.2 Lysing of CHO Cells ...............................................................................................153

4.2.3 Affinity Removal of Expressed Monoclonal Antibody from Lysates .........................154

4.2.4 SDS-PAGE Analysis of Protein A Purified Cell Lysate Fractions ..............................154

4.2.5 Deglycosylation of Glycoproteins ...............................................................................155
4.2.6 Derivatization of Free Glycans with 12C7 Anthranilic Acid and 13C7 Anthranilic Acid..................................................................................................................................155

4.2.7 HPLC Analysis with Fluorescence Detection of Derivatized Glycans ..............156

4.2.8 Quantitative MALDI-TOF MS of Isotopically-Derivatized Glycans...............156

4.3 Results ...............................................................................................................................157

4.3.1 SDS-PAGE Analysis of Protein A Purified Cell Lysate Fractions...............157

4.3.2 HPLC Analysis of Derivatized Glycans from Cell Lysates...............................160

4.3.3 Quantitative Glycomic Analysis of CHO Cells.....................................................161

4.4 Discussion .........................................................................................................................168

4.5 References .........................................................................................................................168

CHAPTER 5

A GENERAL APPROACH FOR THE PURIFICATION AND QUANTITATIVE GLYCOMIC ANALYSIS OF HUMAN PLASMA .........................................................................................................................171

5.1 Introduction .........................................................................................................................172

5.2 Experimental .....................................................................................................................176

5.2.1 Reagents .........................................................................................................................176
5.2.2 Depletion of Human Plasma by MARS .................................................................177

5.2.3 Enrichment of Glycoproteins by MLAC ...............................................................178

5.2.4 Deglycosylation of Glycoproteins .........................................................................179

5.2.5 Derivatization of Free Glycans with $^{12}\text{C}_7$ Anthranilic Acid and $^{13}\text{C}_7$ Anthranilic Acid ..................................................................................................................................180

5.2.6 HPLC Analysis with Fluorescence Detection of Derivatized Glycans ...............180

5.2.8 Quantitative MALDI-TOF MS of Isotopically-Derivatized Glycans ....................181

5.3 Results ...............................................................................................................................181

5.3.1 Premise and Methodology .....................................................................................181

5.3.2 High Abundant Protein Depletion of Human Plasma with MARS 14P ..............184

5.3.3 Enrichment of Glycoproteins by MLAC ...............................................................186

5.3.4 HPLC Analysis of Derivatized Glycans ...............................................................188

5.3.5 Quantitative Glycomic Analysis of Normal vs. Diseased Plasma .........................190

5.4 Discussion .........................................................................................................................192

5.5 References .........................................................................................................................193

CHAPTER 6

PROSPECTIVE FUTURE APPLICATIONS OF THIS ANALYTICAL METHOD ..........199
LIST OF FIGURES

Figure 1.1-1  Biosynthesis of precursor glycan in the ER .............................................................29

Figure 1.1-2  Maturation of glycan in the Golgi apparatus .............................................................30

Figure 1.3-1  Approaches for glycomic analysis ...........................................................................39

Figure 1.3-2  Glyco-catch methodology ........................................................................................41

Figure 1.3-3  Glycan derivatives developed for glycan arrays (Houseman and Mrksich) ............44

Figure 1.3-4  Results from glycan array analysis (Wang et al) .....................................................45

Figure 1.3-5  Results from glycan array analysis (Fukui et al) .......................................................46

Figure 1.3-6  Results from lectin array analysis (Hirabayashi et al) .............................................52

Figure 1.3-7  Results from lectin array analysis (Koshi et al) .......................................................53

Figure 1.3-8  Qproteome GlycoArray (Rosenfeld et al) ..............................................................54

Figure 1.3-9  Strategy for ratiometric lectin array analysis (Pilibello et al) .................................55

Figure 1.3-10 Results from lectin array analysis (Tateno et al) .....................................................56

Figure 1.4-1  Diagram of Antibody Dependent Cellular Cytotoxicity ...........................................67

Figure 1.5-1  Schematic of “laser ionization time-of-flight” mass spectrometer (Tanaka et al) .................................73
Figure 1.5-2  MALDI ionization........................................................................................................74

Figure 1.5-3  Schematic of electrospray ionization with quadrupole mass spectrometer (Fenn et al) ...............................................................................................................77

Figure 1.5-4  Electrospray ionization.............................................................................................78

Figure 1.6-1  Graphic representation of linear TOF mass analyzer ...............................................80

Figure 1.6-2  Graphic representation of reflector TOF mass analyzer .......................................81

Figure 1.6-3  Graphic representation of quadrupole mass analyzer ............................................82

Figure 1.6-4  Cross-section of quadrupole ion trap .....................................................................83

Figure 1.6-5  Graphic representation of ion cyclotron...................................................................85

Figure 2.3.1-1  Workflow for quantitative analysis of glycosylation change by MALDI-TOF MS .........................................................................................................................112

Figure 2.3.2-1  Quantitative MALDI-TOF MS analysis of glycan standards .............................118

Figure 2.3.2-2  Linearity plot for quantitative MALDI TOF MS analysis of glycan standards ..................119

Figure 2.3.3-1  Quantitative MALDI-TOF MS analysis of glycans isolated from model IgG ....................................................................................................................................122
Figure 2.3.3-2  Linearity plot for quantitative MALDI TOF MS analysis of glycans isolated from model IgG

Figure 2.3.4-1  Representative chromatogram of anthranilic acid labelled glycans from model Fc-fusion protein

Figure 2.3.4-2  Quantitative MALDI-TOF MS analysis of glycans isolated from model Fc-fusion protein

Figure 2.3.4-3  Linearity plot for quantitative MALDI TOF MS analysis of glycans isolated from model Fc-fusion protein

Figure 3.3.1-1  Representative chromatogram of anthranilic acid labelled glycans from model Fc-fusion protein

Figure 3.3.1-2  Quantitative MALDI-TOF MS analysis of glycans isolated from Fc-fusion protein at time point 1 vs. time point 6

Figure 3.3.1-3  Line graph of glycans isolated from Fc-fusion protein at time point 1 vs. all other time points

Figure 4.3.1-1  Non-reducing SDS-PAGE analysis of Protein A fractions from cell lysates
Figure 4.3.2-1 Representative chromatogram of derivatized glycans isolated from cell lysates....................................................................................................................................160

Figure 4.3.3-1 Quantitative MALDI TOF MS analysis of glycans isolated from cell line A at time point 1 vs. time point 4..........................................................................................................................164

Figure 4.3.3-2 Quantitative MALDI TOF MS analysis of glycans isolated from cell line B at time point 1 vs. time point 4..........................................................................................................................167

Figure 5.3.1-1 Workflow for purification and quantitative glycomic analysis of human plasma ........................................................................................................................................183

Figure 5.3.2-1 Example chromatogram of MARS 14P depletion of human plasma (Agilent)........................................................................................................................................184

Figure 5.3.2-2 Representative chromatogram of MARS 14P depletion of normal and diseased human plasma..........................................................................................................................185

Figure 5.3.3-1 Representative chromatogram of MLAC enrichment of human plasma (Kullolli et al) ........................................................................................................................................186

Figure 5.3.3-2 Representative chromatograms of MLAC enrichment of normal and diseased human plasma..........................................................................................................................187
Figure 5.3.4-1  Representative chromatogram derivatized glycans isolated from normal and diseased human plasma

........................................................................................................................................189

Figure 5.3.5-1  Quantitative MALDI-TOF MS analysis of normal vs. diseased human plasma

........................................................................................................................................191
LIST OF TABLES

Table 1.2-1  Congenital disorders of glycosylation (affecting \( N \)-glycosylation) .................. 33

Table 2.3.2-1  Statistical data for quantitative MALDI TOF MS analysis of glycan standards ....................................................................................................................................................... 120

Table 2.3.3-1  Statistical data for quantitative MALDI TOF MS analysis of glycans isolated from model IgG .................................................................................................................................................. 124

Table 2.3.4-1  Statistical data for quantitative MALDI TOF MS analysis of glycans isolated from model Fc-fusion protein .......................................................................................................................... 129

Table 3.3.1-1  Quantitative MALDI TOF MS data for glycans isolated from Fc-fusion protein at time point 1 vs. all other time points ........................................................................................................ 143

Table 4.3.3-1  Quantitative glycomics data for cell line A .................................................................................................................. 163

Table 4.3.3-2  Quantitative glycomics data for cell line B .................................................................................................................. 166
## LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbr.</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>Complex glycan, monosialylated biantennary</td>
</tr>
<tr>
<td>A1F</td>
<td>Complex glycan, monosialylated biantennary with core fucose</td>
</tr>
<tr>
<td>A2</td>
<td>Complex glycan, disialylated biantennary</td>
</tr>
<tr>
<td>A2F</td>
<td>Complex glycan, disialylated biantennary with core fucose</td>
</tr>
<tr>
<td>A3</td>
<td>Complex glycan, trisialylated triantennary</td>
</tr>
<tr>
<td>A3F</td>
<td>Complex glycan, trisialylated triantennary with core fucose</td>
</tr>
<tr>
<td>α</td>
<td>alpha</td>
</tr>
<tr>
<td>ADCC</td>
<td>Antibody Depended Cellular Cytotoxicity</td>
</tr>
<tr>
<td>AAL</td>
<td><em>Aleuria aurantia</em> lectin</td>
</tr>
<tr>
<td>β</td>
<td>beta</td>
</tr>
<tr>
<td>12C</td>
<td>Carbon 12</td>
</tr>
<tr>
<td>13C</td>
<td>Carbon 13</td>
</tr>
<tr>
<td>CDC</td>
<td>Complement Dependent Cytotoxicity</td>
</tr>
<tr>
<td>CDG</td>
<td>Congenital Diseases of Glycosylation</td>
</tr>
<tr>
<td>CE</td>
<td>Capillary electrophoresis</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>CHO cells</td>
<td>Chinese Hamster Ovary cells</td>
</tr>
<tr>
<td>ConA</td>
<td><em>Canavalia ensiformis</em> lectin</td>
</tr>
<tr>
<td>Da</td>
<td>Dalton</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethylsulfoxide</td>
</tr>
<tr>
<td>Dol</td>
<td>Dolichol</td>
</tr>
<tr>
<td>ECL</td>
<td><em>Erythrina cristagalli</em> lectin</td>
</tr>
<tr>
<td>EPO</td>
<td>Erythropoietin</td>
</tr>
<tr>
<td>ER</td>
<td>Endoplasmic reticulum</td>
</tr>
<tr>
<td>ESI</td>
<td>Electrospray ionization</td>
</tr>
<tr>
<td>FTICR</td>
<td>Fourier Transform Ion Cyclotron Resonance</td>
</tr>
<tr>
<td>Fuc</td>
<td>Fucose</td>
</tr>
<tr>
<td>γ</td>
<td>gamma</td>
</tr>
<tr>
<td>Gal</td>
<td>Galactose</td>
</tr>
<tr>
<td>GDP</td>
<td>Guanosine diphosphate</td>
</tr>
<tr>
<td>GlcNAc</td>
<td><em>N</em>-acetylglucosamine</td>
</tr>
<tr>
<td>HILIC</td>
<td>Hydrophilic interaction chromatography</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>HPAEC</td>
<td>High pH Anion-exchange chromatography</td>
</tr>
<tr>
<td>IgG</td>
<td>Immunoglobulin G</td>
</tr>
<tr>
<td>HPLC</td>
<td>High Performance Liquid Chromatography</td>
</tr>
<tr>
<td>IFN</td>
<td>Interferon</td>
</tr>
<tr>
<td>IT</td>
<td>Ion trap</td>
</tr>
<tr>
<td>LC-MS</td>
<td>Liquid Chromatograph-Mass Spectrometry</td>
</tr>
<tr>
<td>LIF</td>
<td>Laser Induced Fluorescence</td>
</tr>
<tr>
<td>μ</td>
<td>micro</td>
</tr>
<tr>
<td>μg</td>
<td>microgram</td>
</tr>
<tr>
<td>μL</td>
<td>microliter</td>
</tr>
<tr>
<td>μM</td>
<td>micromolar</td>
</tr>
<tr>
<td>m</td>
<td>milli</td>
</tr>
<tr>
<td>M</td>
<td>Molar</td>
</tr>
<tr>
<td>mA</td>
<td>milliampere</td>
</tr>
<tr>
<td>mAb</td>
<td>monoclonal antibody</td>
</tr>
<tr>
<td>mL</td>
<td>milliliter</td>
</tr>
</tbody>
</table>
mM  millimolar

MALDI  Matrix Assisted Laser Desorption/Ionization

MAL-II  Maackia amurensis lectin

MARS  Multiple Affinity Removal System

MLAC  Multiple Lectin Affinity Chromatography

MS  Mass Spectrometry

MS/MS  Tandem Mass Spectrometry; 2-dimensional Mass Spectrometry

n  nano

NK  Natural Killer cells

nL  nanoliter

nM  nanomolar

NP  Normal Phase

p  pico

P  Phosphate

PAD  Pulsed Amperometric Detection

PEG  Polyethylene glycol
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>PGC</td>
<td>Porous Graphitized Carbon</td>
</tr>
<tr>
<td>pmol</td>
<td>picomol</td>
</tr>
<tr>
<td>PVDF</td>
<td>Polyvinylidene fluoride</td>
</tr>
<tr>
<td>Q</td>
<td>Quadrupole</td>
</tr>
<tr>
<td>RP</td>
<td>Reversed-phase</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodium dodecyl sulfate-Polyacrylamide Gel Electrophoresis</td>
</tr>
<tr>
<td>SNA</td>
<td><em>Sambucus nigra</em> lectin</td>
</tr>
<tr>
<td>SPE</td>
<td>Solid Phase Extraction</td>
</tr>
<tr>
<td>TOF</td>
<td>Time-of-Flight</td>
</tr>
<tr>
<td>UV</td>
<td>Ultra-Violet</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION AND BACKGROUND: GLYCOSYLATION, ITS IMPORTANCE IN BIOLOGY AND BIOTECHNOLOGY, THE EMERGENCE AND EVOLUTION OF GLYCOMICS

1.1 Glycosylation and Glycosynthesis

Asparagine-linked (N-linked) glycosylation, the decoration of oligosaccharide moieties on specific asparagine residues of proteins, was first described in 1961 (Johansen et al 1961). Since then protein glycosylation has been recognized as being the most complicated co- or post-translation modification (Spiro 1973). Accordingly, protein glycosylation has been found to play a multitude of critical roles in biology. These roles range from ensuring correct protein folding and maintenance of solubility and stability (Helenius 2001; Wang et al 1996) to cellular signaling and trafficking (Boscher et al 2011; Walzel et al 2006) to the prevention of proteolysis and infection (Kundra & Kornfeld 1999; Rudd et al 2001; Rudd et al 1999; Sperandio et al 2009; van Veen et al 2004). Glycosylation’s role in proper protein folding and consequently solubility and stability is not completely understood, but many studies have shown that it is required for the proper folding of nascent polypeptides in the endoplasmic reticulum (ER) and for the maintenance of proper folding of mature proteins. Without glycosylation, newly synthesized proteins in the ER soon aggregate and become bound by chaperones that transport them for degradation (Helenius 2001). Meanwhile, the glycosylation of mature proteins have been shown to provide stability against heat stress and positively affects refolding after denaturation (Wang et al 1996). Glycosylation’s roles in signaling and trafficking are highlighted by the involvements of galectins and selectins in various signaling and trafficking events. Galectins are
a family of lectins found in animals, 15 have been identified in mammals, that share a conserved carbohydrate reactive domain (Boscher et al 2011). The binding of galectins to their carbohydrate ligands elicit intracellular signaling events including cell differentiation, proliferation, survival and migration, which in turn affects numerous biological processes, such as embryogenesis, angiogenesis, neurogenesis and immunity (Laderach et al 2010). Selectins are Ca\textsuperscript{2+}-dependent lectins that recognize and bind to specific selectin-ligands; three selectins (P-, E- and L-) are expressed in mammals (Sperandio et al 2009). The transport of leukocytes (leukocyte rolling) during inflammation and infection is, in large part, through their binding of P-selectin glycoprotein ligand-I, which is a heavily glycosylated protein located on the surface of the leukocytes, to P- and E-selectin. Proper glycosylation of selectin ligands is generally required for immune cell trafficking. Glycosylation’s role in preventing proteolysis is not well understood, but numerous experiments have clearly shown this. In the experiments performed by Kundra and Kornfeld, lysozomal-associated membrane proteins I and II (LAMP I and II) were deglycosylated \textit{in vivo} and found to undergo rapid intracellular proteolytic degradation (Kundra & Kornfeld 1999). In the experiments performed by Loh and Gainer, it was observed that the non-glycosylated form of pro-opiocortin (a pro-hormone) was rapidly degraded to peptides within one minute when treated with trypsin, whereas the glycosylated form resulted in two peptides of 23,000 Da and 21,000 Da when treated with trypsin (Loh & Gainer 1980). Glycosylation’s role in the prevention of infection, is rather broad as nearly all of the key players in the immune response are glycoproteins (Rudd et al 2001; Rudd et al 1999). An example can be found with the binding of antibodies to antigen presenting cells, which eventually results in cell lysis by either complement-dependent cytotoxicity or antibody-dependent cellular cytotoxicity. With glycosylation being such an integral part of human biology, it is worthwhile
to describe the process by which glycosylation occurs. The synthesis of N-glycans and their subsequent attachment to proteins occurs in the ER and Golgi complex (Varki et al 2009). The first step in glycosynthesis, occurring on the cytoplasmic side of the ER, is the transfer of a phosphorylated N-acetylglucosamine residue (GlcNAc-P) from the nucleotide-sugar unit uridine diphosphate N-acetylglucosamine (UDP-GlcNAc) to the glycan precursor dolichol phosphate (Dol-P) to yield dolichol pyrophosphate N-acetylglucosamine (Dol-P-P-GlcNAc). The Dol-P-P-GlcNAc is further extended with the addition of another GlcNAc and five mannoses; it is then translocated into the lumen of the ER. Within the ER another four mannoses and three glucose residues are added in a series of enzymatic reactions to yield the Dol-P-P-GlcNAc$_2$Man$_9$Glc$_3$ glycan. This glycan is transferred onto the asparagine residue of the N-glycosylation consensus sequence, A-X-S/T (X is any amino acid other than proline), on various proteins. The glycan, now attached to a protein, is processed by glucosidases I and II and mannosidase I to remove the three glucose residues and one mannose residue resulting a Man$_8$GlcNAc$_2$ glycan. The resulting glycoprotein is then transferred to cis-golgi where it will be processed further by mannosidase I to remove another three mannoses yielding a Man$_5$GlcNAc$_2$ glycan. The glycoprotein is then transferred into the medial-golgi, where the GlcNAc transferase I enzyme (GlcNAcT-1) attaches a GlcNAc residue to the alpha-1,3 linked mannose of the glycan. This step initiates the trimming of the other two extending mannoses by mannosidase II; the resulting glycan can then be acted upon by the GlcNAc transferase II enzyme (GlcNAcT-2) to add a GlcNAc to the alpha-1,6 linked mannose. This Man$_3$GlcNAc$_4$ glycan is the base for all complex-type glycans; from this structure, branching by the addition of GlcNAc residues to the core mannoses can occur. The glycoprotein is then transferred to the trans-golgi where fucosylation, by fucosyltransferase, of the GlcNAc residue attached to the asparagine can occur. The linear extension of the glycan with
galactose by galactosyltransferase, GlcNAc by GlcNAc transferase, and sialic acid by sialyltransferase also takes place in the \textit{trans}-golgi. See Figures 1.1-1 and 1.1-2 for graphical details.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure111.png}
\caption{Synthesis of the dolichol pyrophosphate Glc$_3$Man$_9$GlcNAc$_2$ glycan in the endoplasmic reticulum (ER). Synthesis is initiated on the cytoplasmic side; once the dolichol phosphate is extended by the addition of a phosphorylated GlcNAc and five mannoses, it is translocated into the lumen of the ER. Inside the ER, the four more mannoses and three glucose are added to yield the mature Glc$_3$Man$_9$GlcNAc$_2$ glycan that is transferred to proteins destined for glycosylation. N-acetylglucosamine. Mannose. Glucose. Figure adapted from (Varki et al 2009).}
\end{figure}
Figure 1.1-2. Maturation process of N-glycans. The Glc₃Man₉GlcNAc₂ glycan, now attached to a protein, is transported to the cis-golgi where the three glucose and one mannose are removed to yield the Man₅GlcNAc₂ glycan. The newly trimmed glycoprotein is transferred to the medial-golgi where further processing to remove another five mannoses and add two GlcNAc residues occurs resulting in base Man₃GlcNAc₄ glycan; further additions of GlcNAc to the core mannoses, branching, can also occur. This glycoprotein is then transferred to the trans-golgi where fucosylation and linear extension of the branches takes place. The N-acetylglucosamine.

Mannose. Glucose. Galactose. Figure adapted from (Varki et al 2009).
Because of the integral involvement of glycosylation, it is no surprise that the incorrect glycosylation of proteins or a breakdown in the glycosynthetic pathway would result in some loss of function or other disorder.

### 1.2 Glycosylation and Disease

The biosynthetic pathway of glycans previously described in section 1.1 is that of healthy individuals. Diseases in which the glycosynthetic pathway has been damaged due to mutations in genes coding for various enzymes involved in glycosynthesis have been termed congenital disorders of glycosylation (CDG) (Aebi et al 1999). They have been separated into two groups, type I and type II, with subgroups distinguished by lower case letters (e.g. type Ia, type Ib, etc.) corresponding to the chronological identification of the defective gene. These diseases are all autosomal recessive and hallmarked by a general inability to thrive (Ohtsubo & Marth 2006) with a mortality rate of 20% within the first five years of life (Freeze 2006). Since the first description of CDG in 1980 (Jaeken & Matthijs 2007) a total of forty-five CDG have been identified (Jaeken 2010), with nineteen disorders affecting N-glycosylation (Freeze 2006). Of these nineteen, twelve are type I and seven are type II. All twelve of these type I CDG display hypo- or under glycosylated proteins; this is traced to the insufficient supply and/or incomplete synthesis of the Dol-P-P-GlcNAc$_2$Man$_9$Glc$_3$ glycan. Whichever the aggravating factor, the result for all twelve type I CDG is the expression of proteins with unoccupied glycosylation sites and/or incorrect glycosylation leading to some loss of function. The most common CDG is type Ia, which is known to affect more than 700 patients (Jaeken 2010). The underlying causes are mutations (90 mutations have been identified) to the *PMM2* gene that codes for one of the two phosphomannosmutase enzymes responsible for the conversion of mannose-6-phosphate to mannose-1-phosphate. Mannose-1-phosphate is the precursor for GDP-mannose, which donates
mannose residues for attachment to the growing Dol-P-P-GlcNAc₂Man₅ glycan. This disorder is commonly characterized by inverted nipples, unusual subcutaneous fat deposits, axial hypotonia, internal strabismus (crossed-eyes) and severe psychomotor impairment (Marquardt & Denecke 2003). The remaining type I CDG are far less common than type Ia with less than 100 known patients in total. Type II CDG are identified as diseases causing defects or alterations in the processing of the glycans attached to proteins that have been successfully transferred to the golgi apparatus (Freeze 2006). The resulting symptoms for type II CDG-affected patients varies widely and without commonality; very few are known to be affected by these diseases. Only one patient, each, has been identified as having type IIb and type IIId; the patient affected with type IIb died at 2.5 months. See Table 1.2-1 for a list of type I and type II CDG affecting N-glycosylation with the corresponding defective genes, affected enzymes and characteristic symptoms; the table was adapted from a publication by Hudson H. Freeze (Freeze 2006).
**Table 1.2-1.** Table of congenital disorders of glycosylation (CDG). Table contains information pertaining to many of the identified CDG along with the corresponded defective enzyme and characteristic symptoms displayed by affected patients. Table is adapted from Freeze et al 2006.

<table>
<thead>
<tr>
<th>Disorder</th>
<th>Gene</th>
<th>Enzyme</th>
<th>Characteristic Symptoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDG-Ia</td>
<td>PMM2</td>
<td>Phosphomannomutase II</td>
<td>Mental retardation, hypotonia, esotropia, lipodystrophy, cerebellar hypoplasia, stroke-like episodes, seizures</td>
</tr>
<tr>
<td>CDG-Ib</td>
<td>MPI</td>
<td>Phosphomannose isomerase</td>
<td>Hepatic fibrosis, protein-losing enteropathy, coagulopathy, hypoglycaemia</td>
</tr>
<tr>
<td>CDG-Ic</td>
<td>ALG6</td>
<td>Dol-P-Man: Man₅-GlcNAc₂-P-P-Dol glucosyltransferase</td>
<td>Moderate mental retardation, hypotonia, esotropia, epilepsy</td>
</tr>
<tr>
<td>CDG-Id</td>
<td>ALG3</td>
<td>Dol-P-Man: Man₅-GlcNAc₂-P-Dol mannosyltransferase</td>
<td>Profound psychomotor delay, optic atrophy, acquired microcephaly, iris colobomas, hypsarrhythmia</td>
</tr>
<tr>
<td>CDG-Ie</td>
<td>DPM1</td>
<td>Dol-P-Man synthase I GDP-Man: Dol-P-mannosyltransferase</td>
<td>Severe mental retardation, epilepsy, hypotonia, mild dysmorphism, coagulopathy</td>
</tr>
<tr>
<td>CDG-If</td>
<td>MPDU1</td>
<td>Man-P-Dol utilization 1/Lec35</td>
<td>Short stature, ichthyosis, psychomotor retardation, pigmentary retinopathy</td>
</tr>
<tr>
<td>CDG-Ig</td>
<td>ALG12</td>
<td>Dol-P-Man: Man₇-GlcNAc₂P-P-Dol</td>
<td>Hypotonia, facial dysmorphism, psychomotor retardation, acquired microcephaly, frequent infections</td>
</tr>
<tr>
<td>CDG-Ih</td>
<td>ALG8</td>
<td>Glucosyltransferase II Dol-P-Glc: Man₁-GlcNAc₂-P-P-Dol glucosyltransferase</td>
<td>Hepatomegaly, protein-losing enteropathy, renal failure, hypoalbuminaemia, oedema, ascites</td>
</tr>
<tr>
<td>CDG-li</td>
<td>ALG2</td>
<td>Mannosyltransferase II GDP-Man: Man₁-GlcNAc₂-P-P-Dol mannosyltransferase</td>
<td>Normal at birth; mental retardation, hypomyelination, intractable seizures, iris colobomas, hepatomegaly, coagulopathy</td>
</tr>
<tr>
<td>CDG-Ij</td>
<td>DPAGT1</td>
<td>UDP-GlcNAc: Dol-P-GlcNAc-P transferase</td>
<td>Severe mental retardation, hypotonia, seizures, microcephaly, exotropia</td>
</tr>
<tr>
<td>CDG-Ik</td>
<td>ALG1</td>
<td>Mannosyltransferase I GDP-Man: GlcNAc₂-P-P-Dol mannosyltransferase</td>
<td>Severe psychomotor retardation, hypotonia, acquired microcephaly, intractable seizures, fever, coagulopathy, nephrotic syndrome, early death</td>
</tr>
<tr>
<td>CDG-II</td>
<td>ALG9</td>
<td>Mannosyltransferase Dol-P-Man: Man₆- and Man₈-GlcNAc₂-P-P-Dol mannosyltransferase</td>
<td>Severe microcephaly, hypotonia, seizures, hepatomegaly</td>
</tr>
<tr>
<td>--------</td>
<td>-------</td>
<td>---------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------</td>
</tr>
<tr>
<td>CDG-IIa</td>
<td>MGAT2</td>
<td>GlcNAc transferase 2</td>
<td>Mental retardation, dysmorphism, stereotypies, seizures</td>
</tr>
<tr>
<td>CDG-IIb</td>
<td>GLS1</td>
<td>Glucosidase I</td>
<td>Dysmorphism, hypotonia, seizures, hepatomegaly, hepatic fibrosis; death at 2.5 months</td>
</tr>
<tr>
<td>CDG-IIc</td>
<td>SLC35C1/ FUCT1</td>
<td>GDP-fucose transporter</td>
<td>Recurrent infections, persistent neutrophilia, mental retardation, microcephaly, hypotonia; normal transferrin</td>
</tr>
<tr>
<td>CDG-IId</td>
<td>B4GALT1</td>
<td>β1,4 galactosyltransferase</td>
<td>Hypotonia (myopathy), spontaneous haemorrhage, Dandy– Walker malformation</td>
</tr>
<tr>
<td>CDG-IIe</td>
<td>COG7</td>
<td>Conserved oligomeric Golgi complex subunit 7</td>
<td>Fatal in early infancy; dysmorphism, hypotonia, intractable seizures, hepatomegaly, progressive jaundice, recurrent infections, cardiac failure</td>
</tr>
<tr>
<td>CDG-IIf</td>
<td>SLC35A1</td>
<td>CMP-sialic acid transporter</td>
<td>Thrombocytopenia, no neurological symptoms; normal transferrin, abnormal platelet glycoproteins</td>
</tr>
<tr>
<td>CDG-II/COG1</td>
<td>COG1</td>
<td>Conserved oligomeric Golgi complex subunit 1</td>
<td>Hypotonia, growth retardation, progressive microcephaly, hepatosplenomegaly, mild mental retardation</td>
</tr>
</tbody>
</table>
Diseases due to glycosylation defects other than those classified as CDG also exist. An example and likely the best understood is 1-cell disease (also known as Leroy disease for its discoverer, Jules Leroy), which was the first to be identified as a disease of glycosylation (Leroy & Demars 1967). 1-cell disease is the result of a defective GlcNAc phosphotransferase in the cis-golgi, which results in the cell’s inability to produce mannose-6-phosphate-“tagged” lysosomal enzymes (Durand & Seta 2000). Without the mannose-6-phosphate tag, the lysozomal enzymes are not recognized by the mannose-6-phosphate receptor for transport into the lysosome, which causes a deficiency of multiple enzymes in the fibroblasts and macrophages. Diseases of glycosylation that result in the hyperglycosylation of proteins have also been identified. One of these diseases is Marfan syndrome in which a non-native N-glycosylation site is added to fibrillin 1 (Freeze 2006; Lonnqvist et al 1996; Raghunath et al 1995). The additional glycosylation leads to the inability of the fibrillin monomers to correctly associate. Another disease, which has yet to be named, was reported to be observed in seven children (Vogt et al 2005). These children displayed an extremely high susceptibility to mycobacterial infections. The cause was traced to a defect in the inteferon-gamma receptor, IFNγR2, that caused either a complete loss of binding to INFγ or a substantial loss such that therapy with injections of INFγ was ineffective. The defective IFNγR2 was found to have an additional glycosylation site due to a site mutation at residue 168 from threonine to asparagine, T168N; the glycosylation site was confirmed to be glycosylated by comparing the molecular weights of native and PNGase F-treated wild-type and T168N mutant. Other gains of glycosylation have been observed in diseases affecting coagulation factors, such as Haemaphaelia A; a total of eighteen mutations (missense) resulting in gain of glycosylation have been identified. Another set of diseases not classified as CDG that results in the alteration of glycosylation are those considered acquired diseases. These diseases
are not necessarily congenital, inherited or genetic in nature although some are. Included in this group is a wide range of diseases such as various cancers (Dall'olio 1996), Alzheimer’s disease (Botella-Lopez 2006; Charlwood 2001; Kanninen 2004; Liu et al 2002), rheumatoid arthritis (Kratz et al 2009; Matsumoto et al 2000; Wang et al 2011), diabetes (Carlsson et al 2008; Itoh et al 2007), schizophrenia (Bauer et al 2010; Stanta et al 2010) and liver diseases (Blomme et al 2009; Comunale et al 2010). The alterations in glycosylation varies greatly between diseases. It is these differences that may allow scientists and researchers to identify glyco-biomarkers. To that point, many diseases do display hallmark or characteristic modifications to the glycan structures. For example, many cancers have been observed to have increased levels of sialylation. Colon cancer has been repeatedly reported to present with increased sialylation, specifically α-2,6 linked (Dall'Olio et al 1991; Dall'Olio et al 1992a; b; Saitoh et al 1992); one report presented that approximately 90% of the samples of colon carcinomas tested displayed increased α-2,6 linked sialylation (Dall'Olio et al 1989). Breast (Kyselova et al 2007), ovarian (Saldova et al 2008) and cervical cancers (López-Morales et al 2010) have all been observed to display increased sialylation as well. Meanwhile, rheumatoid arthritis is typically characterized by a decrease in IgG galactosylation (Kratz et al 2009; Matsumoto et al 2000); this has been observed to the extent that it has been recognized as a marker, not only for the presence of rheumatoid arthritis, but also for the progression of the disease (Arnold et al 2007). Difficulties in assigning attributable glycosylation changes may arise when diseases display differences in glycosylation changes that are dependent on sample origin. An example of this can be observed with schizophrenia, in which it was observed that the tetrasialylated tetra-antennary with lactosamine repeat glycan was present in greater abundance in serum, while sialylation was
simultaneously observed to be generally down-regulated in cerebrospinal fluid (Stanta et al 2010).

The previously discussed diseases are only a few examples of how glycosylation and glycosylation changes are associated with human health and disease. To better understand these associations, the field of glycomics was developed.

1.3 Glycomics

More than a decade before the terms “glycome” and “glycomics” emerged there was the term “glycobiology,” which was coined in 1988 by Raymond Dwek and his colleagues of the University of Oxford with the intent of drawing or returning attention to the many integral roles that glycosylation plays in biology (Blow 2009; Rademacher et al 1988). During this time and the many years prior to the coining of the term “glycobiology”, many efforts at better understanding glycosylation were made. Some major examples include: the isolation and characterization of the sugar nucleotides that serve as the substrates in glycosylation, which were originally described in the early 1950’s (Caputto et al 1950; Park 1952); the observation, made in the late 1960’s into the 1970’s, that glycosylation is altered in infected cells, which served as the basis for investigating glycosylation changes in cancer and other diseases (Buck et al 1970; Meezan et al 1969; Van Nest & Grimes 1977); the elucidation of the biosynthetic pathway for N-glycosylation in 1985 (Kornfeld & Kornfeld 1985); the successful commercialization of the first CHO-expressed recombinant glycoprotein, recombinant tissue plasminogen activator, in 1987; all of the roles that glycosylation plays in protein folding, stability, trafficking, transport, cellular signaling, and immunity (Boscher et al 2011; Helenius 2001; Huet et al 2003; Ohtsubo & Marth
The idea of the “glycome” was first proposed in 2000 and 2001 as the “entire set of glycans within an organism” (Hirabayashi et al 2001; Hirabayashi & Kasai 2000). Accordingly, glycomics can be defined or described as the analysis and characterization of all of the glycosylation within a biological system. Due to the extremely complex nature of the glycome, it was understood that a single method to analyze the entire glycome was simply not feasible. Because of this, glycomics has been divided into three branches of study, which has led to the development of numerous methods with varied approaches that aim to provide some insight into the glycome (Varki et al 2009). One branch has focused on the analysis of isolated glycoproteins (and often glycopeptides) to investigate their glycosylation and changes in glycosylation as they relate to disease states; this is essentially a hybridization of proteomics and glycomics and is often called “glycoproteomics”. Another branch has focused on the development of methods targeting the glycosylation on the cell surface in an attempt to understand how glycosylation and changes in glycosylation affect intercellular activity; this can be viewed as a “top-down” approach to glycomics or cellular glycomics. The third branch has focused on the analysis of the glycans isolated from their protein or lipid counter parts; this approach is often sub-divided into the analysis of the N-glycome, O-glycome, glycolipids and glycosaminoglycans, separately. The goal of this approach is to determine what glycan entities are present within a given system and how changes in the expression of these glycans relate to diseases and disease states. A schematic representation of the various approaches is presented in Figure 1.3-1.
Figure 1.3-1. Representation of three main glycomic approaches. (A) describes the analysis of glycans on the surfaces of the cell; these methods generally use antibodies or lectins to probe the glycosylation on the cells. (B) describes the analysis of glycoproteins; the glycoproteins are often digested with proteases to generate glycopeptides; these methods employ mass spectrometric analysis along with antibodies and lectins. (C) describes the analysis of the glycans from glycoproteins; glycoproteins are enzymatically or chemically deglycosylated to yield free glycans; these methods generally employ some type of derivatization of the glycans followed by mass spectrometry. Figure adapted from (Varki et al 2009).
A glycoproteomic method was the first to represent glycomics in its infancy. A method termed “glyco-catch,” was presented by Hirabayashi (Hirabayashi et al 2001; Hirabayashi et al 2002; Hirabayashi & Kasai 2000); the method focused on the elucidation of the glycome of *C. elegans*. In this method, lectins were used to specifically capture glycoproteins by lectin affinity chromatography; the isolated glycoproteins would then be digested with the lysine-specific *Achromaobacter* protease I to generate peptides and glycopeptides; the resulting mixture of peptides and glycopeptides would then be separated over the same previously-used lectin affinity column to specifically capture the glycopeptides; the captured glycopeptides would then be further separated by reversed-phase HPLC. Finally, the glycopeptides would be sequenced and their identities would be determined by database matching/searching; additionally, mass spectrometry could be performed to determine the mass of the attached glycan by subtracting the mass of the peptide from the observed molecular weight. The intent of the method, as expressed by the authors, was to provide a template or “core strategy” for glycomic analysis. See Figure 1.3-2.
Figure 1.3-2. Schematic of workflow for “Glyco-catch” method. The first step of the method utilizes lectin affinity chromatography to isolate glycoproteins. The glycoproteins are then digested with a lysyl peptidase. The resulting glycopeptides are isolated by lectin affinity chromatography. The isolated glycopeptides are separated by reversed-phase HPLC. The purified glycopeptides are then sequenced and identified by database matching/searching. Figure adapted from (Hirabayashi et al 2002).
On the heels of “glyco-catch” came the development of another glycoproteomic method, glycan arrays. This methodology focused on the investigation of glycan-protein interactions. Numerous separate groups (at least six) were involved in its development. In 2002, a number of methods were published on the fabrication of glycan arrays and their use in glycomic analysis (Bryan et al 2002; Fazio et al 2002; Fukui et al 2002; Houseman & Mrksich 2002; Park & Shin 2002; Wang et al 2002; Willats et al 2002). In the method presented by Houseman and Mrksich, the authors described the immobilization of ten monosaccharides to gold-plated glass slides. The method for the production of these “carbochips”, as termed by the authors, required multiple steps to prepare the plates as well as the glycans, and thus the method seems to be quite cumbersome. The plates were modified to have a self-assembling alkanethiol monolayer by submersion into a mixture of two alkanethiols (one with a hydroquinone substitution and one with a penta-ethylene glycol substitution); the resulting monolayer was then oxidized to yield benzoquinone groups, which would serve as the platform for the immobilization of diene-conjugated monosaccharides. The diene-conjugated monosaccharides were synthesized by the coupling of a compound containing a cyclopentene and carboxylic acid functional groups with aminoalkylglycoside derivatives. See Figure 1.3-3. In the method presented by Wang and colleagues, the authors were able to successfully immobilize a series of polysaccharides, different dextrins, onto a nitrocellulose-coated glass slide without a chemical conjugation reaction. These immobilized polysaccharides were shown to retain their antigenicity, as antibodies with known antigen-binding specificities correctly bound to their respective ligands. See Figure 1.3-4. In the method presented by Fukui and colleagues, the authors were able to accomplish the non-covalent immobilization of oligosaccharides onto nitrocellulose and polyvinylidene fluoride (PVDF) membranes; the authors reported better signal with nitrocellulose. The oligosaccharides required a conjugation to
lipid linkers via their reducing ends by reductive amination in order to allow for non-covalent binding to the membranes, as the hydrophilic nature of oligosaccharides prevents this. Ligand (glycan) binding was confirmed by probing with various carbohydrate-binding proteins including antibodies, INF-γ and the chemokine regulated upon activation, normally T-cell expressed and secreted (RANTES). See Figure 1.3-5. The results from all of these reports indicated the feasibility of immobilizing glycans to create glycan arrays to probe the carbohydrate-protein interactions of various proteins within a biological system.

The major drawback for all glycan arrays is the fabrication of the chips, which is often an involved process requiring multiple steps; the major limitation is the availability of glycans to immobilize.
Figure 1.3-3. Compounds used for development of glycan arrays by Houseman and Mrksich. Compound 11 contains a cyclopentene group required for immobilization to the plate through the benzoquinone groups on the plate; it also contains a carboxylic acid group for the conjugation of the aminoalkylglycoside derivatives of the monosaccharides (compounds 1-10). Figure adapted from (Mrksich 2004).
Figure 1.3-4. Figure of glycan arrays generated by Wang and colleagues; the authors were able to successfully immobilize polysaccharides onto nitrocellulose-coated glass slides. The immobilized polysaccharides, dextrans, were probed with antibodies of known antigen specificity to confirm the retention of antigenicity. The bright spots are confirmed binding events. Figure adapted from (Wang et al 2002).
Figure 1.3-5. Figure of glycan arrays generated by Fukui and colleagues; the authors were able to non-covalently immobilize lipid-linked oligosaccharide derivatives on to nitrocellulose membranes. The oligosaccharides were probed with a variety of antibodies, IFN-γ and the chemokine RANTES to confirm that the oligosaccharides maintained their binding properties. The bright spots are confirmed binding events. Figure adapted from (Fukui et al 2002).
Other glycoproteomic methods have been developed that, indeed, utilize the principles of “glyco-catch” by employing some sort of lectin affinity chromatography to specifically enrich for glycoproteins/glycopeptides as part of sample preparation. One of the earlier HPLC-based methods developed for the analysis of glycoproteins was isotope-coded glycosylation-site-specific tagging (IGOT) (Kaji et al 2003). In this method, the authors utilized lectin affinity chromatography in the same manner as in “glyco-catch”. The glycoproteins were enriched from a complex biological sample using an immobilized lectin (agarose-bound concanavalin A, ConA) followed by enzymatic digestion with trypsin of the glycoproteins to generate a mixture of glycopeptides and peptides. The glycopeptides were then isolated from the mixture by lectin affinity with the previously-used lectin column. This method deviated from “glyco-catch” in its subsequent steps. Following enrichment, the glycopeptides were treated with PNGase F in either $^{16}$O- or $^{18}$O-water; upon deglycosylation, oxygen from the water became incorporated into the resulting conversion of the N-linked asparagine to aspartate. The deglycosylated glycopeptides were then analyzed by two-dimensional chromatography in-line with tandem mass spectrometry (2D LC-MS/MS), using anion-exchange as the first chromatographic separation followed by reversed-phase chromatography; mass analysis was performed using a quadrupole-TOF mass spectrometer. This method allowed for the identification of 250 glycoproteins and 400 glycosylation sites in *C. elegans*. A method published in 2004 by Yang and Hancock, introduced the use of multi-lectin affinity chromatography (MLAC) (Yang & Hancock 2005). The authors presented the use of a column composed of a mixture of three lectins: Con A, *Triticum vulgaris* lectin (wheat germ agglutinin, WGA) and *Artocarpus integrifolia* lectin (Jacalin, JAC). These three lectins were chosen to expand the specificity of lectin affinity chromatography used in the enrichment of glycoprotein from human serum. In this method, MLAC was used as the first step
to specifically capture a broad range of glycoproteins from whole human serum; the enriched glycoproteins were then digested with trypsin and the resulting peptides were analyzed by reversed-phase LC-MS/MS. By performing this method, the authors were able to effectively identify 50 glycoproteins in human serum and also determine that approximately 10% of proteins in human serum are glycosylated. Another HPLC-based method published in 2005, utilized serial lectin affinity chromatography (SLAC) - originally developed by Cummings and Kornfeld (Cummings & Kornfeld 1982) - combined with isotope coding and reversed-phase LC-MS/MS to investigate the glycoheterogeneity (difference in glycans occupying a glycosylation site) of glycoproteins in human serum (Qiu & Regnier 2005). In this method, the authors performed enzymatic digestions with trypsin on whole/unfractionated human serum; the resulting peptides were then separated using SLAC with *Sambucus Niger* agglutinin (SNA) and Con A. The first lectin column, immobilized SNA, was used to specifically capture glycoproteins containing α-2,6 linked sialic acid; these glycoproteins were then separated over the immobilized Con A column. The authors collected both the flowthrough and the eluate fractions from the Con A column; the flowthrough was anticipated to contain tri- and tetra-antennary glycans, while the eluate would contain biantennary glycans. The flowthrough fraction was then labelled with normal N-acetoxysuccinimide; the eluate fraction was labelled with D3-N-acetoxysuccinimide. The isotopically-labelled glycopeptides were then combined and analyzed by RPLC-ESI-MS/MS. Mass analysis was performed using a quadrupole-TOF mass spectrometer; by comparing the MS ratios of the isotopically-labelled glycopeptides the authors were able to determine whether the peptide was glycosylated with sialylated biantennary or sialylated tri- and tetra-antennary glycans and to what degree. Tandem mass spectrometry allowed the authors to identify more than 30 of the glycoproteins present in the fractions.
This general workflow of utilizing lectin affinity chromatography to enrich for glycoproteins combined with enzymatic digestions and LC-MS has become standard for HPLC-based glycoproteomic methods. Unfortunately, these methods do not attempt to determine the actual glycosylation of the glycoproteins they identify. Additionally, these methods do not provide quantitative data regarding protein expression.

At around the time, another array-based technology, lectin arrays, was being developed. The concept for lectin arrays was first mentioned in 2004 in a publication by Jun Hirabayashi (Hirabayashi 2004); the purpose of these arrays would be to investigate the lectin-oligosaccharide interaction with the goal of characterizing protein glycosylation. As a means for detection, Hirabayashi suggested the use of evanescent-field fluorescence. In the following year, a method utilizing lectin arrays with evanescent-field fluorescence detection was published by Kuno and colleagues (Kuno et al 2005). In the method, the authors used an array of 39 lectins of known affinities, immobilized on an epoxy-coated glass slide, to demonstrate the ability to profile and differentiate between the glycosylation of various proteins. See Figure 1.3-6. Accordingly, in subsequent years other groups developed their own versions of lectin arrays using different methods of immobilization and detection. In 2006, a method was developed that utilized a supramolecular hydrogel matrix (self-assembled matrix of a low molecular weight organic compound and water) to generate the lectin arrays and a bimolecular fluorescence quenching and recovery (BFQR) method for detection (Koshi et al 2006). The authors reported on the successful use of their method by displaying the correct recognition of simple saccharides (glucose and mannose chains) and model glycoproteins (ribonuclease B, fetuin, ovalbumin) and by characterizing the glycosylation of cell lysates (HepG2 human cells and NM522 bacteria cells). See Figure 1.3.7. In 2007 a method utilizing a prefabricated lectin array chip, in which
the lectins were immobilized onto nitrocellulose, with fluorescence detection was reported (Rosenfeld et al 2007). This reported method was the result of a collaboration between the companies Qiagen and Procognia. The authors reported on the successful glyco-profiling of various glycoproteins including ribonuclease B, prostate specific antigen, porcine thyroglobulin, Tamm horsfall glycoprotein and recombinant human erythropoietin. The highlights of the method were the use of prefabricated lectin arrays and the automated data analysis by the software (Qproteome GlycoArray Analysis) created for the application. See Figure 1.3.8.

Another method published in 2007, reported the use of a 2-dye system to allow for ratiometric quantitation (Pilobello et al 2007). In this method the authors immobilized selected lectins onto prefabricated coated glass slides (Nexterion H slides from Schott, Elmsford, NY). The coating on the slides are described as being a permeable 3-D hydrogel, which sounds similar to the previously mentioned supramolecular hydrogel matrix described by Koshi and colleagues in 2006. Pilobello and colleagues utilized two different fluorescence tags, Cy3 and Cy5, to differentially label their samples. The results of their experiments clearly displayed differential binding of different samples to the lectin arrays. The authors reported on the comparative profiling of cellular-membrane glycosylation of various cells, including 4 different CHO cell lines (normal and 3 mutants: Lec1, Lec2, Lec8) and HL-60 leukemia cells; each of the CHO mutants lack a different glycosyltransferase therefore the glycans expressed by the mutants are markedly different from glycans produced by normal CHO. See Figure 1.3.9. Also in 2007, a method for the characterization of cell surface glycosylation was developed, advancing the lectin array platform to whole-cell glycomics (Tateno et al 2007). The method utilized the same methodology as presented by Kuno et al to immobilize the lectins and detection was achieved by evanescent-field fluorescence. The authors presented the ability to profile and differentiate
between cells based upon the glycosylation presented on the cell surface. The lectin arrays were challenged with CHO cells (normal and 3 mutants: Lec1, Lec2, Lec8), splenocytes (wild-type and a β3GnT2-knockout) and K562 cells (undifferentiated and forced differentiation with sodium butyrate treatment). See Figure 1.3.10. In 2008, Tao and colleagues advanced whole-cell glycomics by probing human cell lines (Tao et al 2008). The authors generated a panel of 94 lectins to investigate 24 human cell lines. The results displayed the ability of the method to differentiate and provide a cell-surface glycan profile for all 24 cell lines. Additionally, experiments were performed to investigate carbohydrate involvement in pathogenic attack on human cells; the results of these experiments indicated the ability of lectin arrays to predict bacterial infection and tropism in human cells. The authors also investigated the ability of the lectin array to function as a tool for biomarker discovery; they were able to show glycosylation differences between cancer stem-like cells grown under different conditions. Ultimately, the authors were able to display the broad applicability of lectin arrays for whole-cell glycomic analysis.

These lectin array methods can perform as either a glycoproteomic method or as a whole-cell glycomic method, depending on the goal of the user; lectin array methods do not disrupt the cell making them an ideal fit for whole-cell glycomics targeted at investigating glycan-mediated cellular interactions. The area of whole-cell or cellular glycomics is, at present, largely under-represented and will likely be further developed in the coming years. As with glycan arrays though, the major drawback of lectin arrays is their fabrication. The availability of lectins, while greater than glycans, is still a limiting factor, as is the lack of truly quantifiable data.
Figure 1.3.6. Lectin arrays of 39 immobilized lectins generated Hirabayashi and colleagues. The lectin arrays were challenged with various glycoproteins. mLam: mouse laminin. ASF: asialofetuin. bTf: bovine transferrin. HRP: horseradish peroxidase. The glycosylation patterns of the glycoproteins were as the authors expected. Figure adapted from (Kuno et al 2005).
**Figure 1.3.7.** (A) Graphic representation of detection by bimolecular fluorescence quenching and recovery on lectin arrays constructed by immobilizing lectins in a supramolecular hydrogel matrix. The fluorescence-labelled lectins are treated with a quenching agent that significantly reduces the fluorescence. Subsequent treatment with an appropriate saccharide ligand restores fluorescence. (B) Lectin arrays composed of six lectins: ConA, AAL, UEA-I, WGA, GSL-II, GSL-I. The arrays were challenged with various simple saccharides, glycoproteins and cell lysates. The slides for ribonuclease B, fetuin, and the lysates of HepG2 and NM522 cells are displayed. Figure adapted from (Koshi et al 2006).
Figure 1.3.8. Graphic representation for the analysis of protein glycosylation using the Qiagen Qproteome GlycoArray slides with Qproteome GlycoArray Analysis software. Two workflows are possible. (A) Direct detection and profiling is achieved by probing the lectin arrays with previously fluorescence-labelled glycoproteins. (B) Indirect detection and profiling is achieved by probing the lectin arrays with glycoproteins then detecting binding with fluorescent-labeled antibodies. Figure adapted from (Rosenfeld et al 2007).
Figure 1.3.9. Graphic representation for the analysis of cellular glycosylation using ratiometric detection with two different fluorescent tags, Cy3 and Cy5. The cells from two different samples are lysed and then centrifuged to isolate the cell membranes. The membranes are then differentially-labelled with either Cy3 or Cy5. The lectin arrays are then probed with an equimolar mixture of the differentially-labelled cell membranes. Figure adapted from (Pilobello et al 2007)
Figure 1.3.10. Results of cell surface glycosylation profiling with lectin arrays; forty-three lectins were immobilized on epoxy-coated glass slides. Normal CHO and three CHO mutants were profiled as whole cells; differential profiles were achieved. Figure adapted from (Tateno et al 2007).
Numerous methods for glycan analysis or glycoprofiling have been developed. One of the earliest chromatography-based methods for the analysis of carbohydrates (Hardy & Townsend 1988) utilized a detection method developed in 1981 by Hughes and Johnson (Hughes & Johnson 1981). In 1981, Hughes and Johnson utilized pulsed amperometric detection (PAD) with high pH (>12-13) to detect 10 simple carbohydrates in solution. Because hydroxyl groups are slightly acidic with pKa values between 12 and 14, they become ionized at high pH; as carbohydrates come in contact with the positively charged electrode, they oxidize creating a current that is detected. In 1988, Hardy and Townsend combined high pH-high pressure anion exchange chromatography with PAD (HPAEC PAD) to analyze oligosaccharides. By taking advantage of the ionization of oligosaccharides at high pH, the authors were able to effectively separate positional isomers of carbohydrates with anion exchange chromatography. In 1993, O'Shea and Lunte published the first method utilizing capillary electrophoresis with PAD to analyze simple carbohydrates (O'Shea & Lunte 1993). In this method, the authors used a polyimide coated capillary with a sodium hydroxide running buffer at pH 12. One of the earliest mass spectrometry methods for carbohydrates utilized fast atom bombardment (FAB) (Dell & Ballou 1983a; Dell & Ballou 1983b; Kamerling et al 1983). In these methods the authors reported on the ability to observe the protonated molecular ion as well as fragment ions which allowed for the determination of the composition and sequence. These three platforms, chromatography, capillary electrophoresis and mass spectrometry, evolved over time as technological advancements were made. Most current methods employ derivatization of the glycans with a fluorophore to allow for highly sensitive detection by fluorescence with chromatography or capillary electrophoresis; derivatization of the glycans also allows for better ionization for mass spectrometry. There are many options of stationary phases available for
modern chromatographic analysis of carbohydrates, including reversed-phase, normal phase/hydrophilic interaction (HILIC), anion exchange and porous graphitized carbon (Pabst & Altmann 2011). The use of reversed-phase is only possible due to the derivatization of the glycans with a fluorescent tag containing a hydrophobic functional group, as glycans only slightly adsorb to the stationary phase on their own. The separation mechanism is based on hydrophobicity; therefore, glycans elute in order of decreasing molecular weight, as larger glycans are more hydrophilic. Normal phase or HILIC columns are generally composed of silica beads modified with amide groups. The separation mechanism is based on the interaction of polar groups with the amide groups; therefore glycans elute in order of increasing molecular weight. Anion exchange, while not necessarily the mode of choice in current methodologies, is still a viable option. The use of high pH allows for the separation of glycans based upon their anionic strength; therefore, glycans elute in order of increasing molecular weight, as larger glycans would contain a greater number of ionized groups. The use of porous graphitized carbon as an analytical HPLC column is a fairly recent development; its use as such for the analysis of carbohydrates was first presented in the early 1990s (Koizumi 1996). The separation mechanism is not fully understood, but it is generally described as being similar to reversed-phase with additional properties; its reversed-phase properties are greater than that of C_{18} and it has the capability to resolve very closely-related and isomeric carbohydrates. Glycans, generally, elute from porous graphitized carbon in order of increasing hydrophobicity. While these chromatographic modes display inherent differences, they are all able to effectively separate oligosaccharides allowing for the profiling of complex mixtures. Capillary electrophoresis of glycans is now often performed utilizing a sieving gel solution that improves separation efficiency (Guttman et al 1996; Szabo et al 2011). Mass spectrometry of oligosaccharides has
largely moved away from fast atom bombardment. Two modes of ionization are now predominantly utilized, electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI), for mass spectrometric analysis of glycans (Rakus & Mahal 2011).

Numerous reports relating to glycomic advances have been published, but it appears that the vast majority has utilized MALDI rather than ESI as the ionization mechanism. While it is unknown exactly why most researchers have chosen MALDI, it is presumed to be due to the advantages afforded with MALDI. The major advantages include less sensitivity to salts, less complex spectra with mainly either +1 or -1 ions, and pmol-fmol sensitivity (Vanderschaeghe et al 2010).

Some examples of glycomic advances utilizing MALDI include the investigation of glycosylation of proteins involved with esophageal cancer (Mechref et al 2009), Alzheimer’s disease (Charlwood 2001), diabetes (Itoh et al 2007), rheumatoid arthritis (Matsumoto et al 2000; Wang et al 2011), prostate cancer (de Leoz et al 2008; Kyselova et al 2007), breast cancer (Kirmiz et al 2007; Kyselova et al 2008), ovarian cancer (An et al 2006), liver diseases (Blomme et al 2009; Comunale et al 2010; Kam et al 2007), colon cancer (Vercoutter-Edouart et al 2008) and stem cells (Satomaa et al 2009). While these reports share the use of MALDI as the ionization source, they differ in detection methods. Some of the authors utilized time-of-flight (TOF), while others utilized Fourier-Transform Ion Cyclotron Resonance (FT-ICR). Some examples of glycomic advances that utilized ESI include the investigation of glycosylation of brain tumors (Prien et al 2008), colon cancer (Korekane et al 2007), colorectal cancer (Misonou et al 2009), ovarian cancer (Kuzmanov et al 2008) and human plasma (Stumpo & Reinhold 2010). Unlike the authors that utilized MALDI, these authors all employed ion traps.
Despite these advances in glycomics that mass spectrometry has provided, it has consistently been criticized for not delivering quantifiable data (Rokus & Mahal 2011). In recent years, some methods have been developed that enable comparative quantitation between samples. In March of 2007, a method utilizing $^{12}$C and $^{13}$C analogs of iodomethane for quantitative glycan analysis was published (Alvarez-Manilla et al 2007). The authors presented data for the qualification/validation experiments that were performed. Initial experiments involved the analysis of oligosaccharides released from human milk; the authors reported data on linearity experiments (5-points) that were performed. Linearity experiments displayed good correlation between the observed and theoretical ratios of $^{12}$C- and $^{13}$C- permethylated glycans with an $R^2$-value of 0.9998. Other proof-of-concept experiments were performed using glycans released from a mixture of three glycoproteins (human alpha 1 acid-glycoprotein, human transferrin and bovine fetuin) to mimic a complex glycomic sample. The mixture of glycoproteins were deglycosylated, isotopically permethylated, and mixed at a 1:1 and 1:3 ratio ($^{12}$C:$^{13}$C); the authors reported an average value of 0.97 with a CV of 12.4% for the 1:1 ratio and an average value of 0.29 with a CV of 16.2% for the 1:3 ratio. With this data, the authors concluded they had successfully developed a method for the quantitative analysis of glycosylation by mass spectrometry. This method was shown to be amenable to both MALDI-TOF MS and ESI-FTICR. Another method utilizing $^{12}$C and $^{13}$C analogs of iodomethane was published in March 2007 (Aoki et al 2007). The isotopic analogs of iodomethane were used to investigate the glycome of the developing *Drosophila* embryo. The authors utilized the isotopic analogs of iodomethane to permethylate released glycans for analysis by, both, MALDI-TOF MS and ESI-IT MS by direct infusion. A brief quality-check of the method was performed by comparing the data of reciprocally-labelled glycans (reversing isotopic analogs of iodomethane for
permethylation reactions of glycans); the results showed comparable values for both labelling procedures. By using isotopic analogs of iodomethane, the authors were able to show differential glycosylation between developmental stages (early and late) of the *Drosophila* embryo. It was specifically mentioned that early-stage embryos presented overall less total glycan mass/abundance and that the glycosylation present was almost entirely high mannose glycans, while late-stage embryos presented with higher total glycan mass/abundance and the glycosylation became more complex with a higher abundance of hybrid and complex-type glycans including some sialylation. With the identification and characterization of glycans from early- and late-stage embryos, the authors ultimately suggested they had essentially accomplished the complete glycomic analysis of the *Drosophila* embryo, with the exception of extremely minor glycans. During the month of August of 2007, another two methods reporting on quantitative analysis by mass spectrometry for glycomics were published. One of these methods was yet another approach utilizing isotopic analogs of iodomethane (Kang et al 2007). This approach was different in that the isotopic analogs were CH$_3$I and CD$_3$I instead of the $^{12}$C and $^{13}$C analogs utilized in the previously discussed methods; also the authors chose to use MALDI-TOF MS, exclusively. The authors performed a linearity experiment (1:6 up to 6:1) with glycans released from ribonuclease B to qualify/validate the method. The authors presented as a real-life application the comparative analysis of glycans from normal sera and cells versus the glycans from the sera and cells of breast cancer patient. The results were not discussed in detail, but the authors did observe an overall increase in glycosylation in the breast cancer samples, which they attribute to aberrant glycosylation often associated with cancer. The other method described the development of four novel isotopic tags for the derivatization of glycans (Bowman & Zaia 2007) and their subsequent analysis by ESI-Q-TOF MS. The authors prepared,
in-house by organic synthesis, 4 isotopic analogs of a derivatization reagent; the tags contained varying amounts of deuterium ranging from 0 up to 12 atoms (0, +4, +8, +12). The tags were appropriate for reductive amination, allowing for a simple and efficient derivatization. Analysis was performed on both chondroitin sulfates A, B, C and glycans released from fetuin labelled with the four tags to demonstrate proof-of-concept; different levels of depolymerized chondroitin sulfate were isotopically labelled and mixed; different amounts of fetuin were deglycosylated and isotopically labelled. The results displayed the appropriate ratios correlating to theoretical values for the mixtures of chondroitin sulfate and fetuin. The authors concluded that the use of their tetraplex label was appropriate for the quantitation of glycosylation changes in complex mixtures. Two methods were published in 2008. The first was published in January; it reported on the development of a method the authors termed “Quantitation by Isobaric Labeling”, QUIBL (Atwood et al 2008). The method entailed the labelling of glycans with isobaric analogs of iodomethane (\(^{13}\text{CH}_3\text{I}\) and \(^{12}\text{CH}_2\text{DI}\)), which yield a mass difference of 0.002922 Da. By using isobaric analogs instead of isotopic analogs, the authors were able to decrease the complexity of the spectra that is generally obtained when analysis is performed with \(^{12}\text{C}\) and \(^{13}\text{C}\) iodomethane. Because the mass difference between glycan pairs increases with increases in glycan size, the mass-differential can become rather large; thereby convoluting the data. The caveat with this isobaric approach is that the mass spectrometer used must be of sufficient resolving capability (on the order of 30,000), as the mass differential between the isobaric labels is quite small; the slight mass differential is somewhat mitigated by the fact that larger glycans will yield larger differences between glycan pairs. A five-point linearity experiment (1:10 up to 10:1) was performed with glycans released from fetuin to qualify/validate the method. The authors were able to apply the method to the analysis of glycans from human serum as a means of providing a
real-life example for proof-of-concept. The method was ultimately applied to differentiating murine embryonic stem cells to show the method’s ability to quantitate glycomic differences; the authors utilized ESI-FTICR MS. The second method of the year was published in July; it reported on the use of $^{12}\text{C}_6$ and $^{13}\text{C}_6$ analogs of aniline to derivatize glycans in a method termed “Oligosaccharide Quantitation using Isotope Tagging” (OliQuIT) by the authors (Ridlova et al 2008). The authors performed a five-point linearity experiment (1:2 up to 10:1) with dextran in order to qualify/validate the method; the results indicated a reasonable linear correlation was achieved. The method was then applied to the oligosaccharides released from plant materials, arabinoxylan from wheat and oat spelt. The differences in the abundances of glycans between the two materials were effectively quantitated. These results displayed the method’s ability to perform quantitative analysis of complex mixtures of glycans; both LC-ESI-TOF MS and LC-MALDI-TOF MS were utilized. In 2009, two more methods aimed at quantitative glycomic analysis were published. The first was published in February. This method also utilized $^{12}\text{C}_6$ and $^{13}\text{C}_6$ analogs of aniline to derivatize glycans; the authors termed this method “Glycan Reductive Isotope Labeling”, GRIL (Xia et al 2009). This methodology was similar to the approach published by Ridlova et al, but these authors investigated the glycans of mammalian systems. Method qualification/validation was accomplished by performing a nine-point linearity experiment (1:10 up to 10:1) with isotopically-labelled glycans from ribonuclease B; $R^2$-values for two of the glycans ($\text{Man}_5$ and $\text{Man}_6$) were reported as 0.998 and 0.988, respectively. The utility of the method was demonstrated by performing analyses on glycans from human and mouse sera. Human sera, from presumably healthy volunteers, were mixed and showed no differences in either the glycans observed or their abundances. In order to highlight the method’s ability to differentiate between glycan abundances in complex glycomic samples, the authors
mixed the glycans from human and mouse sera. The mass spectra showed clear differences in the glycans present and their abundances; MALDI-TOF MS was utilized in this method. The second method of 2009 was published in August. This method utilized an entirely different methodology as compared to the previously discussed methods. The method termed “Isotopic Detection of Amino Sugars with Glutamine”, IDAWG, utilized $^{14}$N and $^{15}$N isotopes of glutamine in cell culture to metabolically incorporate “heavy” nitrogen into amine-containing saccharides (Orlando et al 2009). The authors, with inspiration from Stable Isotope Labeling with Amino acids in Cell culture (SILAC) (Ong et al 2002), utilized the cell’s machinery to generate oligosaccharides containing either $^{14}$N or $^{15}$N. The authors were able to effectively manipulate the cells biosynthetic pathways by understanding that the sole source of nitrogen in glycans comes from glucosamine. It was pointed out the most significant analytical advantage of this method as being the samples can be mixed at the beginning of preparative steps, which minimizes sample handling-related variability. The resulting glycans following cell culture were permethylated prior to mass spectrometric analysis by ESI-FTICR MS. Proof-of-concept was displayed by investigating the glycosylation of proteins expressed by murine embryonic stem cells cultured with either $^{14}$N- or $^{15}$N-glutamine. Expected mass shifts in the glycans were observed (e.g. +2 dalton for high mannose glycans due to the two GlcNAc residues). The authors were able to conclude that this method would be beneficial in the investigation of glycosylation changes or differences occurring during cell culture.

The major drawbacks of most of these methods, particularly those utilizing isotopic permethylation, are the complexity of the data generated and the requirement for very high resolving mass spectrometers. Because permethylation occurs at every available hydroxyl group on glycans, the mass difference due to isotopic derivatization is not consistent between all
glycans; this creates a higher degree of complexity and increases the possibility of overlapping masses. A similar statement can be made regarding the tetraplex method presented by Bowman and Zaia and the metabolic method presented by Orlando and colleagues. The +0, +4, +8 and +12 Da tags of Bowman and Zaia increases the possibility for overlapping masses, as does the series of mass shifts observed with IDAWG due to different amounts of $^{15}$N being incorporated according to the number of N-acetylglucosamines in the glycan.

The importance of glycosylation is not limited to its effect on human health and disease, as it also plays important roles in the therapies that the biopharmaceutical industry develops to help treat many of these diseases.

1.4 Glycosylation and Biotechnology

Within the biotechnology industry, protein glycosylation and its various roles are of high importance, especially considering that the vast majority of commercial and developmental bio-therapeutic products are glycoproteins. Understanding how glycosylation impacts protein folding (Jenkins et al 1996), solubility (Wang et al 1996), stability (Solá & Griebenow 2009), intra- and extra-cellular signaling (Boscher et al 2011; Huet et al 2003; Sperandio et al 2009), receptor binding and effector function and ultimately pharmacokinetics (Li & d’Anjou 2009; Presta 2008) is in large part what allows biotech companies to develop bio-therapeutic drugs. One of the major challenges involved with biotherapeutic development and production is applying this knowledge to the controlling of glycosylation in order to achieve desired glycosylation-related characteristics and properties. The physical properties of potential biotherapeutics are generally the first to be addressed, as proteins that are not physically stable (e.g. improperly folded, low solubility, tendency to aggregate or precipitate, etc.) are not likely to
make it far in the development process. Those that do reach the later stages of development or even commercialization are often investigated for possibilities in improving efficacy and pharmacokinetic properties. One of the primary points of focus is the appropriate binding of the biotherapeutic drug to its intended target. An example where receptor binding was modulated can be observed in experiments involving erythropoietin (EPO) in which the terminal end-capping of glycans with sialic acid (sialylation) was modified; the results of these experiments showed that higher sialylation caused a decrease in receptor binding of EPO (Darling et al 2002; Misaizu et al 1995; Takeuchi et al 1989).

Another major point of focus is the effector function that may be elicited following the binding of a biotherapeutic to its receptor. An example of effector function is antibody-dependent cellular cytotoxicity (ADCC). Antibody dependent cellular cytotoxicity is a mechanism by which the body fights infections (bacterial, virus, etc.). The term “antibody dependent” describes the mechanism by which ADCC occurs; antibodies that recognize cells as being infected or foreign attach to these cells, and in turn natural killer cells (NK) recognize and bind to the antibodies. After the NK-cells bind to the antibody-coated cells, they release enzymes that lead to the apoptosis and cell death (Wallace et al 1994). Many biotherapeutics utilize this mechanism to specifically target diseased cells, as is the case with Rituximab (Maloney et al 1997). Rituximab has been engineered to recognize cancerous B-cells presenting CD-20, and in so doing marks the cell for attack by NK cells; see Figure 1.4-1. It has been observed that ADCC is affected by the glycosylation of the antibody drug product. The observation that core-fucosylation (or the lack thereof) of the glycans greatly increases ADCC activity by increasing antibody binding to the FcγRIII receptor by 50-fold (Niwa et al 2004; Shields 2002; Shinkawa et al 2002), elicited efforts from others to control antibody glycan fucosylation. A genetic approach
in which the gene encoding the fucosyltransferase enzyme, \textit{FUT8}, was knocked out in Chinese hamster ovary (CHO) cells out has been reported (Yamane-Ohnuki et al 2004). By knocking out the gene that codes for the fucosyltransferase enzyme, the scientists were able to achieve the expression of antibodies with glycosylation that is completely devoid of fucose. A chemical approach in which the mannosidase-inhibiting compound, kifunensine, was added to the bioreactor has also been reported (Zhou et al 2008). The use of kifunensine prevents the processing of high-mannose glycans down to the tri-mannosyl chitobiose core glycan, which in turn, prevents the glycosynthesis of complex-type biantennary glycans that generally become fucosylated, resulting in glycosylation composed entirely of non-fucosylated high-mannose structures.

\textbf{Figure 1.4-1.} Antibody dependent cellular cytotoxicity with Rituximab. Rituximab, represented by the conventional “Y-shape”, in blue, recognizes and binds to the B lymphocyte presenting CD20 on its surface. The Fc-receptor on natural killer (NK) cell binds to the Fc-region of Rituximab to initiate cell death of the B lymphocyte. This figure is adapted from www.immunopaedia.org.
Another important attribute is that of serum half-life. The aforementioned case involving EPO can also be used as an example of altering glycosylation as a means of improving this pharmacokinetic property. The decrease in receptor binding in response to increased sialylation was outweighed by the increase in serum half-life that is afforded with increased sialylation. This was highlighted in the case of darbepoetin alfa (rhuEPO) where two additional glycosylation sites were engineered onto the protein and sialylation was increased, resulting in a greater than four-fold increase in serum half-life (Elliott et al 2004). Increasing sialylation results in an increase in serum half-life because the biotherapeutic is protected from clearance by the asialoglycoprotein receptors in the liver; these receptors specifically bind to exposed galactose residues on glycans and removes the attached glycoproteins from circulation.

Additional efforts at controlling glycosylation have been related to controlling bioreactor conditions and cell culture media composition and supplementation. Numerous examples have been reported, ranging from the effects of bioreactor temperature and pH to the inclusion of metal ions to the supplementation with nucleotide-sugar precursors. An example of temperature effects on glycosylation can be found with EPO-Fc in which decreases in terminal sialylation by 20 and 40% were observed when the temperature was reduced to 33 and 30 °C, respectively from 37 °C (Trummer et al 2006). An example of pH affecting glycosylation can be found with EPO in which optimal sialylation levels were achieved at particular pHs; the optimal pH range for EPO was determined to be 6.8-7.2. The potential effect of metal ions can be observed with the addition of manganese to the culture of cells expressing rhuEPO, which resulted in an increase of galactosylation and ultimately an increase in overall glycosylation (Crowell et al 2007). The potential effects of supplementation with nucleotide sugars can be observed in the addition of glucosamine to the cell culture of EPO (Yang & Butler 2002) and the addition of $N$-
acetylmannosamine to the cell culture of IFN-γ (Gu & Wang 1998). The addition of

glucosamine to the bioreactor for EPO resulted in not only a decrease in the sialylation of tetra-

antennary glycans by 41% but also a decrease in the abundance of tetra-antennary glycans by

37%, while the addition of \(N\)-acetylmannosamine to the bioreactor for IFN-γ resulted in the

increase in sialylation of biantennary glycans by 15%.

Within the discussion of controlling the glycosylation of biotherapeutics, it must be mentioned

that of utmost importance to biopharmaceutical companies and regulatory agencies is the

potential for immunologic responses due to glycosylation. Because recombinant biotherapeutics

are often expressed in non-human mammalian cell lines, the glycosylation of the expressed

product may contain non-human glycoforms that may be recognized by the human body’s

immune system triggering an immune response to the drug. A significant example of this is

displayed with cetuximab, which is a chimeric mouse-human monoclonal antibody expressed in

mouse myeloma cells, SP2/0. It was observed that 22% of a particular patient population

suffered from anaphylaxis following administration of the drug (Chung et al 2008). The

underlying causes were determined to be a combination of non-human glycosylation present on

cetuximab and a pre-existing condition in the affected patients. The non-human glycosylation

was determined to be the terminal di-saccharide galactose alpha-1,3 linked galactose (gal \(\alpha\)-1,3

gal). The pre-existing condition was determined to be a hypersensitivity to the terminal gal \(\alpha\)-1,3

gal, which was the result of a previous exposure to an allergen.

Numerous analytical methods were used to characterize the glycosylation of these various

biotherapeutics. Most modern techniques for glyco-analysis utilize either HPLC, capillary

electrophoresis or mass spectrometry (Beck et al 2008). In recent years, mass spectrometry

methods have become increasingly more important and commonplace in industry (Damen et al
Mass spectrometry has played a critical role in the advancement of glycomic analysis. As previously mentioned in section 1.3, the numerous methods developed have utilized two main sources for ionization, MALDI and ESI. Both of these ionization methods are referred to as “soft”, as they allow for the formation of gaseous phase ions with minimal fragmentation to the analyte (Perkel 2001).

The development of MALDI is attributed to Hillenkamp and Karas (Karas et al 1985), as this was the seminal publication from which all subsequent developments with MALDI derive. In this publication, the authors presented the analysis of numerous amino acids and dipeptides to investigate the phenomena of laser desorption. It was observed that the mass spectra of the protonated intact molecule \((M + H)^+\) were achieved with this technique; thereby, presenting the first method to yield intact molecular ions of amino acids. Two wavelengths, 266 and 355 nm, were thoroughly experimented and discussed; it was observed that analysis at 355 nm of dipeptides failed to produce molecular ions. The authors also presented the notion of “matrix-assisted laser desorption” for the first time. An experiment in which tryptophan and alanine were mixed was described; the protonated molecular ion of alanine was observed along with that of tryptophan, but at an irradiation one-tenth the power required to observe alanine alone. It was determined that tryptophan acted as an “absorbing matrix” that transfers energy to the non-absorbing alanine to yield ionization of alanine. Ultimately, the authors concluded that with the
use of an appropriate laser wavelength and a strongly absorbing matrix the deposition of energy to the analyte is more controlled; therefore, allowing for a “soft” ionization that yields molecular ion formation. A few years later in 1988, Tanaka and colleagues published a method utilizing laser ionization with time-of-flight detection (Tanaka et al 1988). In this publication, the authors described the construction of the first laser ionization time-of-flight mass spectrometer; a nitrogen laser (337 nm) was utilized and two modes of analysis, “linear” and “reflection”, were developed in conjunction with a microchannel plate detector that was designed to improve sensitivity. The authors also described using an ultra-fine metal powder to mix with the samples; the metal powder was mixed with glycerol and dissolved in organic solvents and then applied to the sample. By mixing the ultra-fine metal powder with the sample, the authors were able to observe molecular ions for large biologic samples, such as lysozyme (14,306 Da) and chymotrypsinogen (25,717 Da), and the polymers of PEG-4K and PEG-20K; it was reported that without the metal powder it was not possible to detect these ions. This marked the first time that large biologic molecules of this size were successfully detected with mass spectrometry; thereby, allowing the authors to enthusiastically speculate that detection of up to 100,000 Da (m/z) should be feasible with time-of-flight mass spectrometry. This method is what ultimately came to be known as MALDI-TOF MS; the principle author, Koichi Tanaka was awarded the Nobel Prize in Chemistry in 2002 for this work.

The typical MALDI experiment involves the mixing of sample with a matrix at a 1000-fold excess or greater; the matrix used should be UV absorbing near 337 nm to coincide with the nitrogen laser. Multiple strategies exist for depositing (spotting) the sample on to the MALDI target plate; a few examples include the dried-droplet method, vacuum-drying, and the sandwich method. The dried-droplet method is performed by mixing the analyte and matrix solution
followed by spotting onto the target plate; the spot is then allowed to air-dry. The vacuum drying method involves the use of a vacuum chamber to rapidly dry the sample spots; vacuum drying allows for a more even crystallization of the spot to yield more shot-to-shot reproducibility. The sandwich method is performed by spotting the matrix onto the target plate and allowing it to air dry; a spot of analyte is applied directly on top of the dried matrix and allowed to air dry; a second spot of matrix is then applied, resulting in a layer of sample sandwiched between two layers of matrix. Once the samples have been spotted onto the target plate, the plate is introduced into the mass spectrometer where the spots are subjected to quick pulses (up to 1000 shots per second with the current state-of-the art instruments) of the laser. The energy from the laser is transferred through the matrix into to the sample to yield the formation of ions. The ions are then directed into the mass analyzer by controlled electrostatic forces.
1.5-1. Schematic representation of the first “laser ionization time-of-flight” mass spectrometer constructed by Tanaka and colleagues. The authors introduced the “reflective” mode for analysis to improve resolution. A live video camera was also included to allow for monitoring of the target during analysis. Figure adapted from (Tanaka et al 1988).
1.5-2. Graphic representation of MALDI. A laser is pulsed onto a spot of sample and matrix. The laser energy generates a plume/cloud of ions. The ions are directed into the mass analyzer for mass spectrometric detection. Figure adapted from www.nobelprize.org.
The initial development of ESI is attributed to Dole and colleagues in 1968 (Dole et al 1968). In this publication, the authors were the first to report the use of electrospray for the ionization of large macromolecules. Two large polymers of polystyrene, at 51,000 and 411,000 molecular weight, were investigated; it was discovered that dilute solutions (0.01% by weight) were required to obtain an appropriate spray. The other key factor was the desolvation time, as evaporation of the droplets within a “fraction of a second” is required to achieve ionization. This was addressed by the pumping of nitrogen gas into the spray evaporation chamber. The authors concluded that while they were able to successfully achieve electrospray ionization of large macromolecules, there was much more to be learned about the spray and the ions produced.

Some fifteen to twenty years later, Fenn and colleagues publish two pivotal articles reporting on the use of electrospray ionization with HPLC (Fenn et al 1989; Whitehouse et al 1985). In the 1985 publication, the coupling of HPLC to ESI was presented. The authors mostly reported on the development of their new electrospray ionization source and how it should seemingly address the coupling of HPLC to MS. Data using the new ESI source was generated on a handful of samples with molecular weights ranging from a few hundred up to 1,400 daltons. The authors also briefly mentioned the initial experiments performed in which they coupled an HPLC to their ESI instrument; they were able to obtain similar mass spectra for a sample, cyclosporin A, they previously investigated without HPLC. In 1989, the article titled “Electrospray Ionization for Mass Spectrometry for Large Biomolecules” was published. In this publication, Fenn and colleagues presented data on the successful electrospray ionization of proteins ranging from 5,000 up to 76,000 daltons (insulin, cytochrome c, lysozyme, myoglobin, alpha-chymotrypsinogen, alcohol dehydrogenase, alpha-amylase, conalbumin); they reported the ability to analyze as little as a few femtomes of material. Additionally, it was observed that
there was no fragmentation of the samples in the mass spectra; extensive charge multiplicity was also observed. This charge multiplicity was viewed as an advantage, as it effectively extends the mass detection range. The authors concluded with some prospectus regarding the applicability of electrospray ionization to a Fourier Transform Ion Cyclotron Resonance (FTICR) mass analyzer, as the charge multiplicity would address the issues observed with FTICR regarding low ion orbiting frequencies of large molecules. These two publications eventually led to the modern-day HPLC-ESI MS methodology that is now commonly used; John Bennett Fenn was awarded the Nobel Prize in Chemistry in 2002 for this work.

The electrospray ionization of a sample occurs by the flowing of an analyte solution through a capillary tip on which a high electric potential is applied. As the solution flows out of the electrically-charged tip, a fine mist of highly charged droplets is formed. Nitrogen gas and heat are applied to accelerate desolvation of the droplets. As the droplets decrease in size, their charge increases and they become multiply-charged. As the charge to droplet size ratio (charge density) increases, the droplets undergo Coulombic explosions, in which the droplets form smaller droplets; repeated cycles of this process eventually leads to the formation of multiply charged molecular ions. The ions are directed by electrostatic forces into the mass analyzer.
1.5-3. Schematic representation of electrospray ionization quadrupole mass spectrometer used by Fenn and colleagues. Figure adapted from (Fenn et al 1989).
Graphic representation of ESI. The voltage applied to the tip creates a “cone” of charged droplets as the solution passes through. The charged droplets undergo continuous desolvation ultimately leading to ion formation. The ions are guided into the mass analyzer for mass spectrometric detection. Figure adapted from www.nobelprize.org.
1.6 Mass Spectrometry (Mass Analyzers)

Numerous mass analyzers have been developed over time. Four in particular have been of importance in the development of methods for the analysis of glycosylation and biologic samples in general. These include time-of-flight (TOF), quadrupole (Q), ion trap (IT), and Fourier-transform ion cyclotron resonance (FTICR).

The TOF mass analyzer separates ions based on their mass-to-charge ratios (m/z) by determining the time it takes for the ions to travel down a field-free “flight tube” to the detector. The ions are introduced into the “flight tube” by an acceleration voltage; because all of the ions generated have the same charge and amount of energy applied to them courtesy of the acceleration voltage, the ions of smaller mass will travel more quickly, while larger ions will travel more slowly (Cotter 1992). The concept for the TOF mass analyzer was initially proposed by William Stephens in 1946 (Stephens 1946). In his proposal, Stephens described that ions travelling down a vacuum tube would separate according to their different mass-to-charge ratios and result in a spectrum of the different sized ions. In a 1955 publication by Wiley and McLaren, a new ion source containing two ion accelerating regions and two ion focusing techniques to increase resolution for TOF mass spectrometry were described (Wiley & McLaren 1955). The first was described as spatial focusing and was utilized to minimize the spatial distance between ions in their initial state prior to acceleration down the flight tube; this would result in sharper peaks as ions of the same size would reach the detector more closely together. The second was described as time-lag focusing to minimize the effect of the newly generated ions initial velocity prior to acceleration down the flight tube. The idea was to allow some of the initial energy of the newly-formed ions to dissipate by holding them in the source; this would result in better resolution as ions of the same size would travel more closely together down the flight tube. Time-lag focusing
is what led to the development of delayed extraction. In 1972, Mamyrin and colleagues introduced the “reflectron” TOF mass analyzer (Mamyrin et al 1973). The reflectron flight tube is constructed by the inclusion of ion mirrors (repulsive grids) towards the end of the flight tube to bend the ions back towards a detector located towards the beginning of the flight tube. The reflectron provided greater resolution by compensating for discrepancies in the kinetic energies of ions of the same size; more energetic ions will travel farther into the electrostatic field of the reflectron than the less energetic ions. This results in ions of the same size striking the detector simultaneously, yielding higher resolution. Additionally, resolution is increased because the time and distance travelled is increased. This technological breakthrough led to the development of modern-day reflectron TOF mass analyzers that can achieve resolutions on the order of tens of thousands.

![Figure 1.6.1. Graphic representation for a linear time-of-flight mass analyzer. Ions are accelerated out of the source into the flight tube where they separate according to their mass-to-charge ratio. The smaller ions will reach the detector before larger ions.](image-url)
Figure 1.6.2. Graphic representation of reflectron time-of-flight mass analyzer. The ions accelerated from the source travel down the flight tube until they encounter the repulsive ion mirrors. The ions are repelled back towards the source where another detector is located.

The quadrupole (Q) mass analyzer is composed of four metal rods arranged in a parallel fashion. Opposing rods are electrically connected, and a radiofrequency (RF) is applied to each pair. A direct current (DC) is then overlaid or superimposed over the quadrupole. By adjusting the voltages applied to the rods, ions of a particular m/z or m/z range can be selected and allowed to pass through the quadrupole; in this way the quadrupole acts as a mass filter (El-Aneed et al 2009). The concept and development of the quadrupole mass analyzer dates to the early 1950’s and is attributed to Wolfgang Paul. It was actually during his attempts at developing an ion trap that he developed the quadrupole. Paul and colleagues were able to successfully develop the quadrupole ion trap during this time as well. The ion trap, as envisioned by Paul, is a three-dimensional trap consisting of two hyperbolic end-capping electrodes and one ring electrode situated in the middle of the two endcaps (Paul 1990). Ion traps function by the application of an RF voltage to the ring electrode that causes the ions around it become trapped; the kinetic energy
of the ions is dampened by the introduction of helium gas, which serves as molecules for the ions to collide and dissipate their energy. This reduction in kinetic energy makes it easier to trap the ions. The continued trapping of the ions is accomplished by the application of a DC voltage across the endcaps. Ejection of the ions is achieved by the gradual ramping of the RF and DC voltages, which cause destabilization of the trapped ions; the ions eject out of the trap through a hole at the exit endcap electrode in order of increasing m/z. Wolfgang Paul and his colleague, Helmut Steinwendel, received the Nobel Prize in Physics for their development of the quadrupole ion trap.

**Figure 1.6-3.** Graphic representation of quadrupole mass analyzer. Ions are accelerated into the quadrupole; oscillating voltages across the quadrupole selectively allows ions to pass through. This is illustrated by the solid line and dotted line, where the solid line represents ions that were selected to pass and dotted line represents ions that were not allowed to pass. Figure adapted from (El-Aneed et al 2009).
Figure 1.6-4. (A) Cross-section of quadrupole ion trap; its orientation is such that it is situated on the inlet endcap electrode. (B) Graphic representation of the ideal quadrupole ion trap. Figures from (March 1997; 2000)

The Fourier-transform ion cyclotron resonance (FTICR) mass analyzer does not directly measure the m/z of any given ion; rather it measures the frequency with which ions cyclotron (circulate) within a magnetic field (Marshall et al 1998). The FTICR utilizes a Penning ion trap to store and cyclotron ions; within the Penning trap, the ions are contained by a constant axial magnetic field and a constant quadruple electric field; the combination of the fields cause the ions to cyclotron within the given space. By oscillating the electric field, the cyclotron radius of the ions are caused to increase; additionally the excitation causes the ions to form packets and move in phase. Also contained within the Penning trap are a pair of detecting plates that records the frequency with which the ions cyclotron; the detecting plates allows for the storing data for of each
occurrence of the ions’ passing. In this way sensitivity, mass accuracy and resolution are greatly enhanced as each ion is measured/recorded multiple times over; mass resolution of hundreds of thousands up to millions and mass accuracy of single-digit parts-per-million (ppm) can be obtained (Glish & Vachet 2003). The conversion of the collected data to m/z is accomplished by applying Fourier transform mathematics, which serve to deconvolute the ion cyclotron resonance data.

The development of ion cyclotron resonance is attributed to Lawrence and Livingston per their publication on “The Production of Light Ions without the Use of High Voltages” (Lawrence & Livingston 1932). In this publication, the authors described the cyclotron movement of an ion trapped within a uniform magnetic field. It was also shown that the radius of orbit could be increased by applying a transverse electric field, and therefore the ion could be excited to a very high kinetic energy without a large electric field. Earnest Orlando Lawrence was awarded the Nobel Prize in Physics in 1939 for his development and application of the ion cyclotron (Comisarow & Marshall 1996). The first FTICR was not developed and reported until 1974, when Comisarow and Marshall presented data on their very first experiments using their newly-constructed instrument (Comisarow & Marshall 1974a; Comisarow & Marshall 1974b). The fundamental basis on which modern-day FTICR instruments are built and operated were provided by these two early publications by Comisarow and Marshall.
Figure 1.6-5. (A) Schematic representation of ion cyclotron resonance trapping mechanism. The ions initially enter the trap and begin to cyclotron; application of an excitation voltage causes the ions to cluster, cyclotron in phase and increase their cyclotron radius. (B) Simplified cartoon representation of signal generation by cyclotroning ion; as the ions pass by the detection plates a current is recorded; the cyclotron data is Fourier transformed to yield mass spectra. Figure adapted from (Schmid et al 2000).

1.7 Overview of MALDI-TOF MS Method for Quantitative Glycosylation and Glycomic Analysis

An analytical method for the quantitative analysis of glycosylation and glycomic changes by MALDI-TOF MS has been developed. The method employs well-established chemistries and commercially-available reagents, allowing for confidence in and ease of performing analysis. The method allows for the simultaneous analysis of neutral and sialylated glycans in a quantitative manner, while delivering a linear dynamic range over two orders of magnitude.
(isotopic ratios of 0.1 to 10.0) with sub-picomolar sensitivity, high accuracy and high reproducibility. The data is provided in a format that allows for monitoring and quantitating the change in abundance of a single glycan, a particular group or type of glycans, or the entire glycan pool, enabling analysis to be as focused or as broad as needed.

Chapter two describes the method development efforts. Glycan standards, a monoclonal antibody and an Fc-fusion protein were utilized to demonstrate proof of concept. All of the glycans were differentially labelled with isotopic analogs of anthranilic acid ($^{12}$C$_7$ and $^{13}$C$_7$) by reductive amination. The labelled glycans were mixed at various ratios from 1:10 up to 10:1, resulting in a linear correlation (R$^2$-values greater than 0.99) spanning two orders of magnitude. All analyses were performed in triplicate in both negative linear and negative reflector modes. Statistical analysis displays high reproducibility (CVs generally less than 10%) and high accuracy (%Accuracy generally greater than 90%). With the successful development of the method, it was then applied to three different areas of interest.

Chapter three describes the application of the MALDI-TOF MS method to the monitoring of glycosylation changes observed on a biopharmaceutical product over the course of a bioreactor campaign. The interest in this application derives from the biopharmaceutical industry’s interest in better understanding and controlling the biopharmaceutical process. As glycosylation plays critical roles in various aspects of biopharmaceutical products, the industry is keenly interested in finding ways to control the glycosylation of their products.

Chapter four describes the application of the MALDI-TOF MS method to the monitoring of glycomic changes occurring in transfected CHO cells expressing a biopharmaceutical product over the course of a bioreactor campaign. Two different CHO cell lines expressing different
target biotherapeutic antibodies were used in these experiments. Again the interest here is derived from the biopharmaceutical industry’s interest in process control. Whether or not CHO host cell protein glycosylation has any impact on cell health, expression or product quality is unknown, as this work has not been previously performed. The data presented here does not attempt to answer this question as cell health parameters, product titer and product quality data was not available for these materials. Instead, this work is presented as evidence that this a viable method with which to investigate (in future experiments) if any correlations do exist between host cell protein glycosylation and product expression and quality.

Chapter five describes the application of the MALDI TOF MS method in a proposed platform methodology for the preparation and quantitative glycomic analysis of diseased plasma versus normal plasma. High abundant protein depletion with the MARS 14P column and MLAC enrichment of glycoproteins are utilized in the methodology. The glycans from the resulting glycoproteins are removed by treatment with PNGase F and then isotopically labelled and analyzed. Differences in glycosylation between the diseased and normal samples are directly quantitated. The interest with this application is the development of a platform method for quantitative glycomic analysis that can be applied to various biologic fluids to aid in the glycomarker and biomarker discovery effort.

1.8 References


Crowell CK, Grampp GE, Rogers GN, Miller J, Scheinman RI. 2007. Amino acid and manganese supplementation modulates the glycosylation state of erythropoietin in a CHO culture system. *Biotechnology and Bioengineering* 96:538-49


Dall'Olio F, Malagolini N, Serafini-Cessi F. 1992b. The expression of soluble and cell-bound alpha 2,6 sialyltransferase in human colonic carcinoma CaCo-2 cells correlates with the degree of enterocytic differentiation. *Biochem Biophys Res Commun* 184:1405-10


Mamyrin BA, Karataev VI, Shmikk DV, Zagulin VA. 1973. The mass-reflectron, a new nonmagnetic time-of-flight mass spectrometer with high resolution. *Soviet Physics. JETP* 37:45-8


Meezan E, Wu HC, Black PH, Robbins PW. 1969. Comparative studies on the carbohydrate-containing membrane components of normal and virus-transformed mouse fibroblasts. II.
Separation of glycoproteins and glycopeptides by sephadex chromatography.

*Biochemistry* 8:2518-24


Presta LG. 2008. Molecular engineering and design of therapeutic antibodies. *Current Opinion in Immunology* 20:460-70


Chapter 2

DEVELOPMENT OF A MALDI-TOF MS METHOD FOR THE QUANTITATIVE ANALYSIS OF GLYCOSYLATION AND GLYCOMIC CHANGES

Contributing Authors: Samnang Tep, Marina Hincapie and William S. Hancock

Accepted for publication in Carbohydrate Research in conjunction with Chapter 3

Glycoprotein materials provided by Biogen Idec, Cambridge, MA
Chapter 2

DEVELOPMENT OF A MALDI-TOF MS METHOD FOR THE QUANTITATIVE ANALYSIS OF GLYCOSYLATION AND GLYCOMIC CHANGES

2.1 Introduction

Protein glycosylation has been investigated for many years, and these efforts have led to the description of its critical involvement in various biological processes, in health and in disease (see Section 1.2). Because it has been observed that changes in glycosylation are often associated with disease, numerous methods have been developed to attempt to understand and quantify these changes (see Section 1.3). Additionally, due to the interests of the biopharmaceutical industry with glycosylation, numerous methods have been developed to analyze and attempt to control the glycosylation of biological drug products (see Section 1.4). These efforts and findings are the inspiration for this work and it is hoped that the work presented here will help to further advance the field of glycomics. The goal of this work was to develop an analytical method for the quantitative analysis of glycosylation change that would be universally applicable to academic and clinical research as well as the biopharmaceutical industry, thereby bridging the gap that is often observed between these areas. In order to achieve this, the method must utilize not only readily-accessible (commercially-available) reagents and well-established chemistries but also easily-accessible instrumentation that has become fairly commonplace in all settings. Additionally, the method must display good data accuracy and reproducibility, high sensitivity and a broad linear dynamic range. Described here is a method that employs the well-established glycan derivatization chemistry of reductive amination with the readily available derivative anthranilic acid (and its $^{13}$C$_7$ analog) combined with matrix-
assisted laser desorption/ionization-time-of-flight mass spectrometry (MALDI-TOF MS). The $^{13}C_7$ analog creates a 7 Da mass difference, which allows for a large window of separation; thereby ensuring excellent resolution between isotopically-labeled glycan pairs. The method displays a linear dynamic range over two orders of magnitude for both neutral and sialylated glycans. Further the method allows for the simultaneous analysis of neutral and sialylated glycans, which streamlines analysis by removing the requirement of using a different matrix and acquisition mode for neutral and sialylated glycans. Additionally, the use of anthranilic acid allows for chromatographic analysis (Ahn et al. 2010; Anumula & Dhume 1998) with high-sensitivity detection by fluorescence.

2.2 Experimental

2.2.1 Reagents

Complex-type $N$-linked glycan standards and N-Glycanase (PNGase F; EC 3.5.1.52) were purchased from Prozyme (Hayward, CA). Six glycan standards were utilized for their common occurrence in mammalian glycoproteins; three of the glycan standards are neutral (non-sialylated: G0F, G1F, G2F) and three of the glycan standards are charged (sialylated: A1F, A2F, A3). Two purified glycoproteins from Biogen Idec (Cambridge, MA), a monoclonal antibody (IgG) and an Fc-fusion, were used as models. Centrifugal filter (10,000 molecular weight cut off) devices were purchased from Millipore (Billerica, MA). Anthranilic acid, dimethyl sulfoxide and sodium cyanoborohydride were purchased from Sigma Aldrich (St. Louis, MO). $^{13}C_7$ anthranilic acid was purchased from IsoSciences (King of Prussia, PA). Acetic acid, acetonitrile and HPLC-grade water were purchased from Fisher Scientific (Pittsburg, PA). Post-derivatization cartridges (LudgerClean S) were purchased from Ludger, Ltd. (Oxfordshire, UK).
Solid phase extraction blocks (Oasis HLB) were purchased from Waters Corporation (Milford, MA).

2.2.2 Derivatization of Free Glycans with $^{12}\text{C}_7$ Anthranilic Acid and $^{13}\text{C}_7$ Anthranilic Acid

Free glycans (either standards or released from glycoproteins) were derivatized by reductive amination with 350 mM anthranilic acid ($^{12}\text{C}_7$ or $^{13}\text{C}_7$), 1 M sodium cyanoborohydride in a 70:30 DMSO:acetic acid solution. The derivatization was allowed to proceed for 3 hours at 45 °C. When glycan standards were used, a single 10 µg aliquot (of each standard) was split into two equal parts prior to derivatization; 5 µg were labeled with standard $^{12}\text{C}_7$ anthranilic acid, while the other 5 µg were labeled with $^{13}\text{C}_7$ anthranilic acid. When either the IgG or Fc-Fusion were used, the glycans were isolated from 100 ug protein; the isolated glycans from each (the IgG and Fc-Fusion) were split into two equal parts prior to derivatization. Half of each were labeled with standard anthranilic acid, while the other half were labeled with $^{13}\text{C}_7$ anthranilic acid. Post-derivatization purification of all labeled glycans was performed using LudgerClean S cartridges following the manufacturer’s suggested protocol. The purified, labeled glycans were evaporated to dryness in a vacuum centrifuge.

2.2.3 Deglycosylation of Glycoproteins

Prior to the deglycosylation reaction, 100 ug of each glycoprotein (IgG and Fc-fusion protein) was buffer exchanged into 50 mM ammonium bicarbonate to remove any potential interfering sample matrix components and to adjust the pH closer to the pH optimum (8.6) of N-Glycanase. Once buffer exchanged, the two glycoproteins were treated with N-Glycanase at 37 °C for 16 to 20 hours (overnight). Following deglycosylation, the released glycans were isolated via reversed phase solid phase extraction using Oasis HLB 96-well blocks. By performing reversed phase
solid phase extraction, the released glycans are collected in the flowthrough, while the deglycosylated protein and enzyme are bound to the stationary phase. This procedure is straightforward as it avoids using a bind/elute protocol that may result in the loss of some glycans if they are not completely eluted from the stationary phase. The flowthroughs were evaporated to dryness in a vacuum centrifuge in preparation for derivatization.

2.2.4 HPLC Analysis with Fluorescence Detection of Derivatized Glycans

The derivatized glycans were analyzed by HPLC with fluorescence detection. Increased sensitivity with fluorescence (ex: 330 nm; em: 420 nm) is afforded with the use of anthranilic acid, thereby allowing for detection of lower abundance glycans and/or the use of less analyte. The glycans are separated over a Shodex AsahiPak NH2-P column using an increasing ammonium acetate gradient. Mobile phase A is 0.5 mM ammonium acetate at pH 5.0; mobile phase B is 250 mM ammonium acetate at pH 5.0; chromatographic separation is achieved using a two-staged linear gradient over 85 minutes. Performing this type of HPLC analysis allows for a previewing of the types of glycans present in the sample. With an understanding of where certain glycans and types of glycans elute using this particular combination of column, mobile phases and gradient, a powerful orthogonal method can be utilized to aid in peak identification.

2.2.5 Quantitative MALDI-TOF MS of Derivatized Glycans

MALDI-TOF MS of the derivatized glycans was performed using an AB-Sciex 5800 TOF-TOF instrument (Foster City, CA). The alternately-derivatized glycan standards and glycans released from the two model glycoproteins were analyzed in both negative linear and negative reflector modes. The evaporated glycan samples were reconstituted in purified water and then mixed at seven ratios (1:10, 1:5, 1:2, 1:1, 2:1, 5:1, 10:1) with their respective isotopically-labeled
counterparts. The matrix utilized for all sample types was anthranilic acid (10 mg/mL) in 50% acetonitrile-water. The glycan mixtures were applied to the MALDI target plate using the dried-droplet-sandwich method. The use of anthranilic acid as the tag and matrix allows for the analysis of both neutral and sialylated glycans, separately and simultaneously. This greatly simplifies the methodology as it avoids the necessity of using two different matrices and two different modes (positive and negative ion) of MS acquisition to perform the analysis of a single complex sample containing both neutral and sialylated glycans.

2.3 Results

2.3.1 Premise and Methodology

The goal of this method is to provide a means of quickly acquiring meaningful quantitative data for changes in glycosylation. By derivatizing free glycans with either a normal (12C7) or heavy (13C7) tag, quantitation between differentially labeled glycan samples can be achieved. In this way, changes to even a single glycan can be observed as the data is presented as glycan pairs separated by 7 Da. Additionally, changes to a particular class or group of glycans can be observed. For instance, changes in overall sialylation can be determined by monitoring the change in ratios for all sialylated glycan pairs between samples. This type of information could be very powerful, as the ability to easily monitor the changes in glycosylation over time would be immensely informative; it could yield information pertaining to glycosylation changes relating to cellular development and differentiation or the progression of diseases. Perhaps, over time, the method could become diagnostic or predictive with continued use. A workflow of the methodology is described in Figure 2.3.1-1. It should be pointed out that in the workflow, the post-labelling purification is performed on the individual samples rather than a mixture of the
differentially-labelled glycans. This may present an additional source of variability, but the results from the method qualification experiments with glycan standards and two model glycoproteins indicate this approach yields very good data accuracy and reproducibility for all glycans tested. Additionally, this was the most facile approach to generate the series of commixes/ratios used for method qualification. It is, however, acknowledged that commixing the samples prior to post-labelling purification should yield even greater data reliability.

Because this method relies on the quantitation of isotopically-labelled glycan pairs, it was extremely important to establish the linear dynamic range of the ratios. This was accomplished by mixing equimolar amounts of the $^{12}\text{C}_7$ and $^{13}\text{C}_7$ anthranilic acid derivatized glycans in a series of ratios (1:10, 1:5, 1:2, 1:1, 2:1, 5:1, 10:1). Individual glycan standards and the mix of released glycans from each of the glycoprotein samples were separately analyzed using these ratios. All analyses were performed in triplicate and in both negative linear and negative reflector modes. The rationale for performing analysis in both linear and reflector modes was to determine whether or not all of the glycans, especially the sialylated glycans, could be effectively analyzed in reflector mode, as it has been reported that analysis of sialylated glycans in reflector mode can lead to the loss of sialic acid and decreased sensitivity (Papac et al 1996). The ability to perform analysis in reflector mode is desired, as it allows for greater resolution.
Figure 2.3.1-1. Workflow for sample preparation and analysis of samples for quantitative glycan analysis by MALDI-TOF MS. The first step is the deglycosylation of glycoprotein samples by treatment with PNGase F to obtain free glycans. The second step is the alternate derivatization of the free glycans with normal $^{12}$C$_7$ anthranilic acid and $^{13}$C$_7$ anthranilic acid and the subsequent post-labelling purification. Finally, the evaporated glycans are reconstituted and mixed at a 1:1 ratio and analyzed by MALDI-TOF MS to quantitate any differences in glycosylation between the samples.
2.3.2 Quantitative MALDI-TOF MS Analysis of Derivatized Glycan Standards

Six glycan standards were selected for their common occurrence in mammalian glycoproteins. Three neutral (G0F, G1F, G2F) and three sialylated (A1F, A2F, A3) were used. The glycan standards were individually analyzed to demonstrate proof of concept. Without demonstrating that quantitative analysis of a single differentially labeled glycan could be achieved, the method would not be suitable for the analysis of biotherapeutic samples, as they typically contain a complex mixture of glycans. Because the glycan standards were singularly pure (90% or greater purity) entities, they provided the least complex system to investigate. Additionally, the only sample prep required prior to derivatization was the splitting of the aliquot into two equal parts. A reproducible linear correlation was established for each glycan standard over the range of 1:10 to 10:1 in both negative linear and negative reflector modes; data were easily acquired with low picomolar quantities. The ratios were calculated based upon the peak intensity of the monoisotopic peak of each derivatized glycan from the accumulation of at least 1000 spectra. The linear response achieved with the sialylated glycans in negative reflector mode was a welcomed result, as it indicated the capability of the method to accurately perform analysis in reflector mode. It should be noted that the resolution of the A3 glycan substantially decreased at the extremes (1:10 and 10:1) of the linear range in negative linear mode; this was not observed in negative reflector mode, as resolution of the glycan present at 10% was consistently greater than 8000. Observed CVs were generally less than 10% at any given point, while data accuracy was generally 90% or greater. See Figures 2.3.2-1, -2 and Table 2.3.2-1.
Figure 2.3.2-1. MALDI-TOF MS spectra of each of the $^{12}\text{C}_7$- and $^{13}\text{C}_7$ anthranilic acid-labelled glycan standards, mixed at a 1:1 ratio, in negative reflector mode. Isotopic resolution of the glycan mass is achieved in reflector mode; the monoisotopic masses are labelled. The accuracy is maintained for all neutral and sialylated glycans. (A) Spectrum of isotopically-labelled G0F. (B) Spectrum of isotopically-labelled G1F. (C) Spectrum of isotopically-labelled G2F. (D) Spectrum of isotopically-labelled A1F. (E) Spectrum of isotopically-labelled A2F. (F) Spectrum of isotopically-labelled A3. Cartoon representations of the glycan structures have been included. ♦Fucose. ■N-acetylglucosamine. ○Mannose. ♣Galactose. ♣Sialic acid (Neu5Ac).
Figure 2.3.2-2. Linearity data for the six glycan standards mixed at seven ratios (1:10, 1:5, 1:2, 1:1, 2:1, 5:1, 10:1); data collected in negative reflector mode. Each glycan standard was analyzed separately.
### Statistical Data for Quantitative MALDI-TOF MS Analysis of Isotopically Labeled Glycan Standards

<table>
<thead>
<tr>
<th>Theoretical</th>
<th>G0F</th>
<th>G1F</th>
<th>G2F</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Observed</td>
<td>RSD</td>
<td>%Accuracy</td>
</tr>
<tr>
<td>1:10 (0.1)</td>
<td>0.11</td>
<td>9.1%</td>
<td>90.0</td>
</tr>
<tr>
<td>1:5 (0.2)</td>
<td>0.23</td>
<td>4.3%</td>
<td>85.0</td>
</tr>
<tr>
<td>1:2 (0.5)</td>
<td>0.56</td>
<td>8.8%</td>
<td>88.0</td>
</tr>
<tr>
<td>1:1 (1.0)</td>
<td>0.96</td>
<td>1.8%</td>
<td>96.0</td>
</tr>
<tr>
<td>2:1 (2.0)</td>
<td>1.87</td>
<td>7.2%</td>
<td>93.5</td>
</tr>
<tr>
<td>5:1 (5.0)</td>
<td>4.7</td>
<td>5.6%</td>
<td>94.0</td>
</tr>
<tr>
<td>10:1 (10.0)</td>
<td>9.14</td>
<td>9.1%</td>
<td>91.4</td>
</tr>
<tr>
<td></td>
<td>A1F</td>
<td></td>
<td>A2F</td>
</tr>
<tr>
<td>1:10 (0.1)</td>
<td>0.11</td>
<td>5.1%</td>
<td>86.7</td>
</tr>
<tr>
<td>1:5 (0.2)</td>
<td>0.23</td>
<td>4.9%</td>
<td>85.0</td>
</tr>
<tr>
<td>1:2 (0.5)</td>
<td>0.55</td>
<td>9.1%</td>
<td>90.0</td>
</tr>
<tr>
<td>1:1 (1.0)</td>
<td>1.01</td>
<td>1.1%</td>
<td>99.0</td>
</tr>
<tr>
<td>2:1 (2.0)</td>
<td>1.71</td>
<td>4.4%</td>
<td>85.7</td>
</tr>
<tr>
<td>5:1 (5.0)</td>
<td>4.77</td>
<td>4.8%</td>
<td>95.4</td>
</tr>
<tr>
<td>10:1 (10.0)</td>
<td>9.44</td>
<td>10.2%</td>
<td>94.4</td>
</tr>
</tbody>
</table>

**Table 2.3.2-1.** Reproducibility and accuracy for data acquired by MALDI-TOF MS in negative reflector mode for all six glycan standards used at all seven ratios. The ratios are presented as $^{12}$C$_7$: $^{13}$C$_7$. The reproducibility is generally less than 10%, and the accuracy is generally greater than 90%.

### 2.3.3 Quantitative MALDI-TOF MS Analysis of Derivatized Glycans from a Model IgG

Having established the method’s capability to perform quantitative analysis of individual differentially labeled glycans, the method was challenged with a more complex sample containing a mixture of glycans. The method was demonstrated to have the ability to simultaneously quantitate a mixture of glycans released from an IgG. Since IgGs generally contain three predominant glycans (G0F, G1F and G2F) that are all neutral, they provided a simplified system with which to examine this point. While other glycans were observed, data analysis was focused on these three glycans. As with the experiments performed on the glycan standards, triplicate analysis at each ratio was performed in both linear and reflector modes.
Understanding that the glycans of IgGs generally constitute approximately 3% of the total mass, it was calculated that the glycans released from the IgG were detected and quantitated at the sub-picomolar level; data were acquired below 0.1 pmol. A linear correlation was achieved with both linear and reflector modes for all three glycans; ratios were calculated based upon the peak intensity of the monoisotopic peak of each derivatized glycan from the accumulation of at least 1000 spectra. Observed CVs were generally less 10% at any given point, while data accuracy was generally 90% or greater. See Figures 2.3.3-1, -2 and Table 2.3.3-1.
Figure 2.3.3-1. MALDI-TOF MS spectra of glycans from the model IgG labelled with isotopic analogs of anthranilic acid. (A) IgG glycans mixed at a 1:10 ratio (\(^{12}C_7:^{13}C_7\)). (B) IgG glycans mixed at 1:1 ratio (\(^{12}C_7:^{13}C_7\)). (C) IgG glycans mixed at 10:1 ratio (\(^{12}C_7:^{13}C_7\)). Cartoon representations of the glycan structures have been included. ▲:Fucose. □:N-acetylgalactosamine. ⊱:Mannose. ○:Galactose. ●:Sialic acid (Neu5Ac).
Figure 2.3.3-2. Linearity data for the glycans isolated from the model IgG mixed at seven ratios (1:10, 1:5, 1:2, 1:1, 2:1, 5:1, 10:1); data collected in negative reflector mode.
Table 2.3.3-1. Reproducibility and accuracy for data acquired by MALDI-TOF MS in negative reflector mode for the glycans from the model IgG. The ratios are presented as $^{12}\text{C}_7$:$^{13}\text{C}_7$. The reproducibility is generally less than 10%, and the accuracy is generally greater than 90%.

2.3.4 Quantitative MALDI-TOF MS Analysis of Derivatized Glycans from a Model Fc-Fusion Protein

In order to demonstrate that the method could effectively provide quantitative data for a complex mixture of neutral and sialylated glycans, an Fc-fusion protein containing both neutral and sialylated glycans was utilized. The Fc-fusion protein was previously characterized by HPLC with fluorescence detection, see Figure 2.3.4-1; the major glycans observed were G0F, G1F, G2F, A1F, and A2F. These five glycans were therefore the focal points for MALDI-TOF MS analysis. Again, as previously performed for the other sample types, triplicate analysis at each ratio in both negative linear and negative reflector modes was performed. Understanding that the glycans of this particular Fc-fusion protein constitute approximately 10% of the total mass, it was calculated that the glycans released from the Fc-fusion protein were detected and quantitated at the sub-picomolar level; data were acquired below 0.3 pmol. A linear correlation was achieved for all glycans (neutral and sialylated) in both modes; ratios were calculated based upon
the peak intensity of the monoisotopic peak of each derivatized glycan from the accumulation of at least 1000 spectra. Observed CVs were generally less 10% at any given point, while data accuracy was generally 90% or greater. See Figures 2.3.4-2, -3 and Table 2.3.4-1.
Figure 2.3.4-1. Representative chromatogram of anthranilic acid-labelled glycans from the model Fc-fusion protein. The major neutral glycans (G0F, G1F, G2F) and sialylated glycans (A1F, A2F) are highlighted. These glycans were selected for monitoring by MALDI-TOF MS analysis. Cartoon representations of the glycan structures have been included. ◼Fucose. □N-acetylglucosamine. ●Mannose. ○Galactose. ◆:Sialic acid (Neu5Ac).
Figure 2.3.4-2. MALDI-TOF MS spectra of glycans from the model Fc-fusion protein labelled with isotopic analogs of anthranilic acid. (A) IgG glycans mixed at a 1:10 ratio ($^{12}$C$_7:^{13}$C$_7$). (B) IgG glycans mixed at 1:1 ratio ($^{12}$C$_7:^{13}$C$_7$). (C) IgG glycans mixed at 10:1 ratio ($^{12}$C$_7:^{13}$C$_7$). Cartoon representations of the glycan structures have been included. ▲: Fucose. □: N-acetylglucosamine. ●: Mannose. ○: Galactose. ◆: Sialic acid (Neu5Ac).
Figure 2.3.4-3. Linearity data for the glycans isolated from the model Fc-fusion protein mixed at seven ratios (1:10, 1:5, 1:2, 1:1, 2:1, 5:1, 10:1); data collected in negative reflector mode.
<table>
<thead>
<tr>
<th>Theoretical</th>
<th>G0F Observed</th>
<th>RSD</th>
<th>%Accuracy</th>
<th>G1F Observed</th>
<th>RSD</th>
<th>%Accuracy</th>
<th>G2F Observed</th>
<th>RSD</th>
<th>%Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1:10 (0.1)</td>
<td>0.11</td>
<td>5.4%</td>
<td>93.3</td>
<td>0.10</td>
<td>5.6%</td>
<td>96.7</td>
<td>0.11</td>
<td>5.4%</td>
<td>93.3</td>
</tr>
<tr>
<td>1:5 (0.2)</td>
<td>0.22</td>
<td>5.3%</td>
<td>91.7</td>
<td>0.22</td>
<td>5.3%</td>
<td>91.7</td>
<td>0.22</td>
<td>5.2%</td>
<td>88.3</td>
</tr>
<tr>
<td>1:2 (0.5)</td>
<td>0.59</td>
<td>1.9%</td>
<td>81.3</td>
<td>0.59</td>
<td>1.0%</td>
<td>81.3</td>
<td>0.59</td>
<td>3.9%</td>
<td>82.7</td>
</tr>
<tr>
<td>1:1 (1.0)</td>
<td>1.01</td>
<td>10.0%</td>
<td>99.0</td>
<td>1.01</td>
<td>9.8%</td>
<td>98.7</td>
<td>1.00</td>
<td>9.0%</td>
<td>99.9</td>
</tr>
<tr>
<td>2:1 (2.0)</td>
<td>2.10</td>
<td>5.0%</td>
<td>95.0</td>
<td>2.21</td>
<td>1.3%</td>
<td>89.5</td>
<td>2.16</td>
<td>1.2%</td>
<td>92.0</td>
</tr>
<tr>
<td>5:1 (5.0)</td>
<td>4.37</td>
<td>7.9%</td>
<td>87.3</td>
<td>5.09</td>
<td>3.0%</td>
<td>98.2</td>
<td>4.48</td>
<td>2.5%</td>
<td>89.6</td>
</tr>
<tr>
<td>10:1 (10.0)</td>
<td>9.1</td>
<td>9.1%</td>
<td>91.4</td>
<td>9.39</td>
<td>5.6%</td>
<td>93.9</td>
<td>9.14</td>
<td>9.1%</td>
<td>91.4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>A1F</th>
<th>G0F Observed</th>
<th>RSD</th>
<th>%Accuracy</th>
<th>G1F Observed</th>
<th>RSD</th>
<th>%Accuracy</th>
<th>G2F Observed</th>
<th>RSD</th>
<th>%Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1:10 (0.1)</td>
<td>0.11</td>
<td>5.4%</td>
<td>93.3</td>
<td>0.11</td>
<td>5.4%</td>
<td>93.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1:5 (0.2)</td>
<td>0.22</td>
<td>5.2%</td>
<td>88.3</td>
<td>0.22</td>
<td>5.2%</td>
<td>88.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1:2 (0.5)</td>
<td>0.59</td>
<td>1.7%</td>
<td>82.0</td>
<td>0.59</td>
<td>3.4%</td>
<td>82.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1:1 (1.0)</td>
<td>1.06</td>
<td>8.8%</td>
<td>94.3</td>
<td>1.03</td>
<td>9.0%</td>
<td>96.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2:1 (2.0)</td>
<td>2.19</td>
<td>1.3%</td>
<td>90.5</td>
<td>2.00</td>
<td>2.0%</td>
<td>100.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5:1 (5.0)</td>
<td>4.48</td>
<td>2.5%</td>
<td>89.6</td>
<td>4.69</td>
<td>2.7%</td>
<td>93.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10:1 (10.0)</td>
<td>9.44</td>
<td>10.2%</td>
<td>94.4</td>
<td>9.14</td>
<td>9.1%</td>
<td>91.4</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2.3.4-1. Reproducibility and accuracy for data acquired by MALDI-TOF MS in negative reflector mode for the glycans from the model IgG. The ratios are presented as $^{12}$C$_7^{13}$C$_7$. The reproducibility is generally less than 10%, and the accuracy is generally greater than 90%.

2.4. Discussion

A MALDI-TOF MS method for the quantitative analysis of protein glycosylation has been described. Reductive amination of free glycans with isotopic analogs of anthranilic acid make it feasible to explore the glycosylation from most any type of glycoprotein sample. The type of data provided, the ratio of isotopically-labelled glycan pairs, allows for the determination of glycosylation changes to a single glycan, a group of glycans, or the entire glycan pool. This kind of flexibility is important as the monitoring of changes to a particular glycan or type(s) of glycan(s) may yield information leading to a better understanding of relation between glycosylation change and development or disease.
Admittedly, similarities exist between this method and other previously developed methods that utilized isotopic analogs of aniline (Ridlova et al 2008; Xia et al 2009). These similarities include the use of reductive amination with tags containing carbon isotopes and the ability to analyze the derivatized glycans by either MALDI-TOF MS or LC-ESI MS (Kurihara et al 2009; Yu et al 2010). The advantages with this methodology come from the use of anthranilic acid as the derivatization reagent, which allows for detection by fluorescence if analyzed chromatographically or electrophoretically (Kamoda et al 2006). The use of fluorescence detection is advantageous versus the other similar methods where UV detection was used with aniline analogs as the derivatization reagents, as the sensitivity of fluorescence detection can be up to 1000 times greater than UV detection. Also, by performing chromatographic or electrophoretic analysis with fluorescence detection, a preview of the glycans that are present as well as the overall relative quantitation of glycans can be achieved with minimal sample consumption; this allows for the preservation of the bulk of sample material for MALDI-TOF MS analysis. Further, the use of anthranilic acid as the tag and matrix allows for the simultaneous analysis of neutral and sialylated glycans. While seemingly slight, this is a significant improvement as it obviates both the need for different matrices for neutral and sialylated glycans as well as the analysis in both positive and negative modes of a single sample containing neutral and sialylated glycans, thereby streamlining analysis. While the method, as presented, does not investigate MS/MS analysis for structural/linkage information, tandem MS analysis by either LC-ESI-MS/MS (Harvey 2005; Wuhrer et al 2009) or MALDI-TOF/TOF (Wuhrer & Deelder 2005) MS, which would yield some structural/linkage information should be feasible.
In the subsequent chapters, this method will be utilized to quantitate glycosylation changes occurring on a purified biopharmaceutical product over the course of a bioreactor campaign, glycomic changes occurring within CHO cells (two cell lines) expressing a biopharmaceutical product, and glycomic changes occurring in human plasma between normal and disease-afflicted individuals.

2.5 References


Kurihara T, Min JZ, Hirata A, Toyo'oka T, Inagaki S. 2009. Rapid analysis of N-linked oligosaccharides in glycoproteins (ovalbumin, ribonuclease B and fetuin) by reversed-


Chapter 3

A MALDI-TOF MS METHOD FOR THE SIMULTANEOUS AND QUANTITATIVE ANALYSIS OF NEUTRAL AND SIALYLATED GLYCANS OF CHO-EXPRESSED GLYCOPROTEINS

Contributing Authors: Samnang Tep, Marina Hincapie and William S. Hancock

Accepted for publication in Carbohydrate Research in conjunction with Chapter 2

Glycoprotein materials provided by Biogen Idec, Cambridge, MA
Chapter 3

A MALDI-TOF MS METHOD FOR THE SIMULTANEOUS AND QUANTITATIVE ANALYSIS OF NEUTRAL AND SIALYLATED GLYCANS OF CHO-EXPRESSED GLYCOPEPTIDES

Accepted for publication in Carbohydrate Research

3.1 Introduction

Within the biopharmaceutical industry there is a continuous drive towards better understanding and controlling glycosylation. This drive is fueled, in large part, by the fact that the vast majority of developmental and commercially-available bio-therapeutic drugs are glycoproteins. Because of the various critical roles that glycosylation plays in biology, it behooves the industry to be keenly aware of glycosylation and how it impacts their products. Of particular interest are properties relating to protein folding (Helenius 2001), solubility (Wang et al 1996) and stability (Solá & Griebenow 2009), as potential products would need to display correct folding, low levels of aggregation and high stability under various conditions. Of even greater importance are the roles that glycosylation plays in cellular signaling (Boscher et al 2011; Huet et al 2003; Sperandio et al 2009) and protein-protein interactions, such as receptor binding and effector function, and pharmacokinetics (Li & d’Anjou 2009; Presta 2008). The glycosylation of bio-therapeutic drugs can directly affect the binding of the drug to its intended target. An example of this is displayed in the case of erythropoietin (EPO) (Misaizu et al 1995; Takeuchi et al 1989). Studies involving the use of three variants of EPO with differing levels of sialylation (terminal sialic acid residues), showed that the more highly sialylated version displayed a three-fold decrease in association with the receptor as compared to the least sialylated version. The
glycosylation of monoclonal antibody drugs has been shown to have a direct impact on antibody dependent cellular cytotoxicity (ADCC), and because of this many efforts have been made to control antibody glycosylation (Jefferis 2009a; Jefferis 2009b; Presta 2008). Particular attention, as of late, has been paid to the fucosylation of the glycan moieties of antibodies, as the lack of fucose has been shown to increase ADCC activity (Niwa et al 2004; Shields 2002; Shinkawa et al 2002). A gene engineering approach, in which the FUT8 gene responsible for coding the enzyme whose function it is to fucosylate glycans was knocked out, was achieved by Yamane-Ohnuki and colleagues (Yamane-Ohnuki et al 2004). A chemical approach with the mannosidase-inhibiting compound kifunensine was utilized to prevent mannosidase activity resulting in non-fucosylated high mannose glycans, was achieved by Zhou and colleagues (Zhou et al 2008). The glycosylation of bio-therapeutics can also affect their pharmacokinetic properties by altering clearance and serum half-life. Glycans with less than complete sialylation or a complete lack of sialylation are quickly targeted for removal by the asialoglycoprotein receptor in the liver, while high-mannose glycans and terminal N-acetylglucosamine containing glycans are targeted by the mannose receptor. An example of increased serum residency time is demonstrated with darbepoetin alfa (rhuEPO) in which two additional glycosylation sites were engineered onto the protein and sialylation was increased (Elliott et al 2004). An example of the effect of terminal N-acetylglucosamine on clearance can be observed with lenecrecept (an Fc-fusion protein), in which it was shown that protein entities bearing exposed glycans with terminal N-acetylglucosamine residues were selectively removed from circulation within the first 12 to 24 hours following administration (Jones et al 2006). The glycosylation-related property of the highest importance to biopharmaceutical companies and regulatory agencies is the potential for immunogenic responses due to glycosylation (Buttel et al 2011). Because recombinant bio-
therapeutics are often expressed in non-human mammalian cell lines, the glycosylation of the expressed product may contain non-human glycoforms that may be recognized by the human body’s immune system triggering an immune response to the drug. A major example of this was observed with cetuximab (a monoclonal antibody) in which 22% of a particular patient population suffered from anaphylaxis following administration of the drug (Chung et al 2008). The underlying causes were determined to be a combination of non-human glycosylation present on Cetuximab and a pre-existing condition in the affected patients. The non-human glycosylation was determined to be the terminal di-saccharide galactose alpha-1,3 linked galactose (gal α-1,3 gal). The pre-existing condition was determined to be a hypersensitivity to the terminal gal α-1,3 gal due to the previous exposure to an allergen.

Numerous methods have been developed over the years to provide glycosylation data, generally in the form of glycosylation profiles. These methods range from chromatographic (Ahn et al 2010; Anumula & Dhume 1998; Dhume et al 2008) to electrophoretic (Kamoda et al 2006; Szabo et al 2011) to mass spectrometric (Damen et al 2009; Lim et al 2008; Mauko et al 2011; Wagner-Rousset et al 2008). While these methods perform extremely well at providing analytical tools to profile and characterize glycosylation, they do not yield quantitative data between samples. Without the ability to quantitate the glycosylation differences between samples, important information may be missed. An example of this would be the scenario in which there is a difference in glyco-occupancy between samples but their glycosylation profiles are the same; in this situation, profiling methods would not be able to detect the difference. Another scenario would be the situation in which certain purification steps specifically remove or enrich (unknowingly) for certain glycans or types of glycans; in this case, standard profiling methods would detect a change in the relative distribution of glycans, but they would not be able
to discern if a particular glycan or set of glycans is actually increasing or decreasing or whether
the changes are due to cell culture conditions or the purification steps. The method discussed in
Chapter 2 was utilized here to quantitatively monitor the changes in glycosylation occurring on a
CHO-expressed bio-therapeutic over the course of a bioreactor campaign; six time points were
analyzed. The bio-therapeutic analyzed was an Fc-fusion protein bearing both neutral and
sialylated glycosylation; colleagues within Biogen Idec performed centrifugation and Protein A
purification to remove cell debris and CHO-host cell proteins.

3.2 Experimental

3.2.1 Reagents

The purified Fc-fusion protein, sampled at six time points, was provided by Biogen Idec.  Centrifugal filter (10,000 molecular weight cut off) devices were purchased from Millipore (Billerica, MA).  N-Glycanase (PNGase F; EC 3.5.1.52) was purchased from Prozyme (Hayward, CA).  Anthranilic acid, dimethyl sulfoxide and sodium cyanoborohydride were purchased from Sigma Aldrich (St. Louis, MO).  $^{13}$C$_7$ anthranilic acid was purchased from IsoSciences (King of Prussia, PA).  Acetic acid, acetonitrile and HPLC-grade water were purchased from Fisher Scientific (Pittsburg, PA).  Post-derivatization cartridges (LudgerClean S) were purchased from Ludger, Ltd. (Oxfordshire, UK).  Solid phase extraction blocks (Oasis HLB) were purchased from Waters Corporation (Milford, MA).

3.2.2 Deglycosylation of the Fc-Fusion Protein

Prior to the deglycosylation reaction, 100 ug of the Fc-fusion protein at each time point was
buffer exchanged into 50 mM ammonium bicarbonate to remove any potential interfering sample
matrix components and to adjust the pH closer to the pH optimum (8.6) of N-Glycanase. Once buffer exchanged, the glycoprotein samples were treated with N-Glycanase at 37 °C for 16 to 20 hours (overnight). Following deglycosylation, the released glycans were isolated via reversed phase solid phase extraction using Oasis HLB 96-well blocks. By performing reversed phase solid phase extraction, the released glycans are collected in the flowthrough, while the deglycosylated protein and enzyme are bound to the stationary phase. This procedure is straightforward as it avoids using a bind/elute protocol that may result in the loss of some glycans if they are not completely eluted from the stationary phase. The flowthroughs were evaporated to dryness in a vacuum centrifuge in preparation for derivatization.

3.2.3 Derivatization of Free Glycans with $^{12}$C$_7$ Anthranilic Acid and $^{13}$C$_7$ Anthranilic Acid

Free glycans released from 100 µg of the Fc-fusion protein at each time point were derivatized by reductive amination with 350 mM anthranilic acid ($^{12}$C$_7$ or $^{13}$C$_7$), 1M sodium cyanoborohydride in a 70:30 DMSO:acetic acid solution. The derivatization was allowed to proceed for 3 hours at 45 °C. Post-derivatization cleanup of all labelled glycans was performed using LudgerClean S cartridges following the manufacturer’s suggested protocol. The purified, labelled glycans were evaporated to dryness in a vacuum centrifuge.

3.2.4 HPLC Analysis with Fluorescence Detection of Derivatized Glycans

The derivatized glycans were analyzed by HPLC with fluorescence detection. Increased sensitivity with fluorescence (ex: 330 nm; em: 420 nm) is afforded with the use of anthranilic acid; thereby, allowing for detection of lower abundance glycans and/or the use of less analyte. The glycans were separated over a Shodex AsahiPak NH2-P column using an increasing ammonium acetate gradient. Mobile phase A is 0.5 mM ammonium acetate at pH 5.0; mobile
phase B is 250 mM ammonium acetate at pH 5.0; chromatographic separation was achieved using a two-staged linear gradient over 85 minutes. Performing this type of HPLC analysis allows for a previewing/profiling of the types of glycans present in the sample, as well as the determination of their relative abundances, if desired. With an understanding of where certain glycans and types of glycans elute using this particular chromatographic method, a powerful orthogonal method that aids in the identification of peaks observed in mass spectrometry is afforded.

3.2.5 Quantitative MALDI-TOF MS of Isotopically Derivatized Glycans

MALDI-TOF MS of the derivatized glycans was performed using an AB-Sciex 5800 TOF-TOF instrument (Foster City, CA). The evaporated glycan samples were reconstituted in purified water and then mixed. The $^{12}\text{C}_7$ anthranilic acid derivatized glycans from the Fc-fusion protein at the first time point were mixed at a 1:1 ratio with the $^{13}\text{C}_7$ anthranilic acid derivatized glycans at each of the successive bioreactor time points and then analyzed so that any changes in glycosylation could be observed and quantitated. The matrix utilized was anthranilic acid (10 mg/mL) in 50% acetonitrile-water. The glycan mixtures were applied to the MALDI target plate using the dried-droplet-sandwich method. Each mixture was analyzed in triplicate, and final spectra were the accumulation of at least 1000 individual spectra.

3.3 Results

3.3.1 Quantitative MALDI-TOF MS Analysis of Derivatized Glycans from an Fc-Fusion Protein Over the Course of a Bioreactor Campaign
The Fc-fusion protein was analyzed at six different time points of a bioreactor run in order to observe and quantitate any changes in its glycosylation that may have occurred. Five glycans were chosen for monitoring based upon HPLC analysis of the derivatized glycans; the three major neutral glycans and two different sialylated glycans were chosen. See Figure 3.3.1-1. The glycans from the first time point were derivatized with $^{12}\text{C}_7$ anthranilic acid, while the glycans from each of the successive time points were labeled with $^{13}\text{C}_7$ anthranilic acid. The glycans from the first time point were mixed at a 1:1 ratio with the glycans from each of the other successive time points; final spectra were the accumulation of at least 1000 individual spectra and isotopic ratios were calculated based upon the intensity of the isotopic peak of each derivatized glycan. This allowed for the observation of any changes in glycosylation that occurred relative to the first time point. The results from the mixture of the first and second time points showed an upward shift in the abundances of all glycans at the second time point; this trend was observed through time point five. At time point six, a decrease in the abundance of all glycans was observed. The neutral glycans (G0F, G1F and G2F) returned to levels similar to those observed at the second time point, while the sialylated glycans (A1F and A2F) decreased to levels below those observed at the initial time point. See Figures 3.3.1-2, -3 and Table 3.3.1-1. This type of decrease in terminal sialylation of glycans may be important, as it could impact the pharmacokinetic properties of this protein. These results could be used to aid in making modifications to either or both the bioreactor conditions and time of harvest in order to optimize for the product quality attribute of terminal sialylation.
Figure 3.3.1-1. Representative chromatogram of anthranilic acid-labelled glycans from the Fc-fusion protein. The major neutral glycans (G0F, G1F, G2F) and two sialylated glycans (A1F, A2F) are highlighted. These glycans were selected for monitoring by MALDI-TOF MS analysis. Cartoon representations of the glycan structures have been included. ✧ Fucose. ✦ N-acetylglucosamine. ◀: Mannose. ○: Galactose. ♦: Sialic acid (Neu5Ac).
Figure 3.3.1-2. MALDI-TOF MS spectrum of $^{12}$C$_7$- and $^{13}$C$_7$ anthranilic acid-labelled glycans, in negative reflector mode, released from 100 µg of an Fc-fusion protein at time point 1 ($^{12}$C$_7$) and time point 6 ($^{13}$C$_7$) of a bioreactor run; the glycans were mixed at a 1:1 ratio to observe any changes in abundance of the glycans. The glycan ratio values are presented as $^{13}$C$_7$: $^{12}$C$_7$ in order to visualize changes relative to the first time point; values greater than 1 indicate an increase in abundance, while values less than 1 indicate a decrease in abundance. Five glycan pairs are highlighted: G0F, G1F, G2F, A1F and A2F; these glycans were previously identified by HPLC analysis. The average relative ratios, presented as time point 6 ($^{13}$C$_7$) vs. time point 1 ($^{12}$C$_7$), of each glycan pair is presented above the highlights. ●: Fucose. □: N-acetylglucosamine. ●: Mannose. ●: Galactose. ●: Sialic acid (Neu5Ac).
Figure 3.3.1-3. Line graph displaying the changes in the ratios between time point 1 and each successive time point of the five highlighted glycan pairs. The glycans show a general upward trend up to time point 5. At time point 6, a decline in all of the glycans is observed; the neutral glycans (G0F, G1F, G2F) return to similar levels as observed at time point 2; the sialylated glycans (A1F, A2F) show a decline to levels below those observed at time point 1. The glycan ratio values are presented as $^{13}C_C^{12}C_C$ in order to visualize changes relative to the first time point; values greater than 1 indicate an increase in abundance, while values less than 1 indicate a decrease in abundance.

Table 3.3.1-1. Quantitative MALDI-TOF MS data for glycans from Fc-fusion protein over the course of a bioreactor campaign. The average ratios between each time point and time point 1 are displayed. The glycan ratio values are presented as $^{13}C_C^{12}C_C$ in order to visualize changes relative to the first time point; values greater than 1 indicate an increase in abundance, while values less than 1 indicate a decrease in abundance.
3.4 Discussion

The changes in glycosylation occurring on a CHO-expressed biotherapeutic over the course of a bioreactor campaign have been quantitated by MALDI-TOF MS; the method utilized isotopic analogs of anthranilic acid to generate isotopic derivatives of glycans. Mass spectrometric analysis of the isotopically-labelled glycans allowed for the determination of product glycosylation differences between time points. The data acquired indicated a decrease in product sialylation at the final time point (harvest date); this piece of information is important, as it could impact the pharmacokinetic properties of the product. This type of information could be used to help design future cell culture conditions. Additionally, this information could be combined with other product quality data to yield a better understanding of how the product is impacted by the various steps in process development. To that point, while this method, as presented here, was only applied to the analysis of glycosylation change as it related to bioreactor incubation time, it would also be appropriate for comparative analysis of glycosylation for cell line screening activities, cell culture development, purification development, and formulations development (the entirety of process development activities).

From this work, came the notion of applying this method towards the investigation of the CHO glycome. With as much effort made by the biopharmaceutical industry to understand and control the biopharmaceutical process, it is a wonder why no efforts have been made to try to understand the glycome of the CHO cells expressing a bio-therapeutic product. The work presented in Chapter 4 will endeavor to fill this gap.

3.5 References


Presta LG. 2008. Molecular engineering and design of therapeutic antibodies. *Current Opinion in Immunology* 20:460-70


IgG1 Complex-type Oligosaccharides Show the Critical Role of Enhancing Antibody-dependent Cellular Cytotoxicity. *Journal of Biological Chemistry* 278:3466-73


producing completely defucosylated antibodies with enhanced antibody-dependent cellular cytotoxicity. *Biotechnology and Bioengineering* 87:614-22

Chapter 4

THE CHARACTERIZATION AND QUANTITATION OF GLYCOMIC CHANGES IN CHO CELLS DURING A BIOREACTOR CAMPAIGN

Contributing Authors: Samnang Tep, Marina Hincapie and William S. Hancock

Submitted for publication in Biotechnology and Bioengineering

Chinese hamster ovary cells were cultured and harvested by colleagues within the Cell Culture Development group at Biogen Idec, Cambridge, MA. A special thanks to Rashmi Korke.
Chapter 4

THE CHARACTERIZATION AND QUANTITATION OF GLYCOMIC CHANGES IN CHO CELLS DURING A BIOREACTOR CAMPAIGN

Submitted for publication in Biotechnology and Bioengineering

4.1 Introduction

Within the biopharmaceutical industry much attention has been paid towards product glycosylation. Rightfully so, as glycosylation plays numerous important roles including protein folding (Helenius 2001), solubility (Wang et al 1996), stability (Solà & Griebenow 2009), cellular signaling (Boscher et al 2011; Huet et al 2003; Sperandio et al 2009) and pharmacokinetics (Li & d’Anjou 2009; Presta 2008). Numerous efforts have been made to control product glycosylation; many of these efforts have focused on the modification of bioreactor conditions and cell culture media composition and supplementation. Reports have been made regarding the effects of bioreactor temperature and pH to the inclusion of metal ions to the supplementation with nucleotide-sugar precursors. Bioreactor temperature was shown to affect the sialylation of an Fc-fusion of erythropoietin (EPO-Fc); EPO-Fc sialylation was observed to decrease by 20% and 40% when the temperature was reduced to 33 ºC and 30 ºC, respectively, from 37 ºC (Trummer et al 2006). Bioreactor pH was shown to affect the sialylation of EPO (Yoon et al 2005); it was observed that the optimal pH range was 6.8-7.2 for optimal sialylation. The addition of manganese (Mn²⁺) to the bioreactor was shown to increase galactosylation and overall glycosylation (glyco-site occupancy) of EPO (Crowell et al 2007). Supplementation with glucosamine was shown to decrease the abundance of tetra-antennary glycans by 37% and decrease tetra-antennary sialylation by 41% on EPO (Yang & Butler 2002).
Meanwhile, supplementation with N-acetylmannosamine was shown to increase biantennary sialylation of interferon-gamma (IFN-γ) by 15% (Gu & Wang 1998).

Interestingly, no efforts have been made to investigate the glycome of the cells that are used to express bio-therapeutics. This comes as a bit of a surprise since the notion that cellular glycosylation changes over time is not new. Because of the lack of efforts made in this area, there are many questions regarding whether and how the glycome of Chinese hamster ovary (CHO) cells affects cellular health, productivity and product quality. Presented here is the application of the method described in Chapter 2 towards the quantitative analysis of glycomic changes occurring in CHO cells over the course of a bioreactor campaign. Two different cell lines will be investigated; each expressing a different bio-therapeutic. The goal of this work is to provide a viable method by which the glycomic analysis of bio-therapeutic expressing cells can be performed and combined with correlative investigations.

4.2 Experimental

4.2.1 Reagents

Rapigest surfactant and Oasis HLB solid phase extraction blocks were purchased from Waters Corporation (Milford, MA). Protein A HPLC column was purchased from Life Technologies/Applied Biosystems (Carlsbad, CA). Bovine serum albumin protein standards and coomassie reagent for the Bradford assay and coomassie stain were purchased from Bio-Rad (Hercules, CA). 4-20% Tris-Glycine pre-cast polyacrylamide gels, molecular weight standards, 2X sample buffer and 10X running buffer were purchased from Life Technologies/Invitrogen (Carlsbad, CA). Porous graphitized carbon solid phase extraction blocks were purchased from Thermo Fisher Scientific (Pittsburg, PA). N-Glycanase (PNGase F; EC 3.5.1.52) was purchased
from Prozyme (Hayward, CA). Cells from two CHO cell lines expressing monoclonal antibodies were provided by Biogen Idec (Cambridge, MA) for glycomic analysis. Anthranilic acid, dimethyl sulfoxide and sodium cyanoborohydride were purchased from Sigma Aldrich (St. Louis, MO). $^{13}$C$_7$ anthranilic acid was purchased from IsoSciences (King of Prussia, PA). Halt protease inhibitor cocktail, Acetic acid, acetonitrile and HPLC-grade water were purchased from Thermo Fisher Scientific (Pittsburgh, PA). Post-derivatization cartridges (LudgerClean S) were purchased from Ludger, Ltd. (Oxfordshire, UK). Centrifugal filtration devices (5,000 molecular weight cut off) were purchased from Millipore (Billerica, MA).

### 4.2.2 Lysing of CHO Cells

Two chinese hamster ovary cell lines were genetically modified to, each, express a different monoclonal antibody. Both cell lines (cell line A and cell line B) were grown under fed-batch conditions in 3-liter sparged bioreactors using a proprietary medium; cell line A was grown for thirteen days and cell line B was grown for 11 days. Cell number and viability were measured using a Cedex automated cell counter (Innovatis, Bielefeld, Germany) that uses trypan blue staining. Cells, equivalent to a volume containing $1\times10^7$ cells, were removed at four different time points of each cell culture/bioreactor campaign and pelleted by centrifugation (1,000 for 2 minutes). The pellets were resuspended in 5 mL of 1X phosphate-buffered saline (PBS) and then centrifuged again (1,000 g for 1 minute); the supernatants were removed and the pellets were frozen and stored at -70 ºC until needed for analysis.

For lysing, the pellets were resuspended in 0.5 mL 1X Tris-buffered saline (TBS) supplemented with 0.1% Rapigest surfactant and 5 µL of Halt protease inhibitor cocktail (Carlage et al 2009). To lyse the cells, the suspensions were subjected to three 15-second cycles of sonication. The
lysates were then centrifuged at 10,000 rpm for 10 minutes to pellet the cell debris. The supernatants, containing intracellular and membrane proteins, were aspirated and retained for analysis. The total protein concentration of each of the lysates was determined by the Bradford assay.

### 4.2.3 Affinity Removal of Expressed Monoclonal Antibody from Lysates

In order to remove the expressed monoclonal antibodies from the lysates, the lysates were passed over a Protein A column. Since Protein A specifically binds to the Fc-region of antibodies, the antibodies in the lysates would be retained on the column, while the rest of the proteins in the lysates would flow through. Both the flow through and eluate fractions were collected; all of the fractions were buffer-exchanged into 50 mM ammonium bicarbonate in preparation for deglycosylation with N-Glycanase. The total protein concentration of each of the fractions was determined by the Bradford assay.

The chromatographic method utilized a four-buffer system, including an equilibration buffer, a salt wash, an elution buffer and a regeneration buffer. The equilibration buffer consists of 25 mM sodium phosphate and 200 mM sodium chloride at pH 7.4. The salt wash consists of 1 M sodium chloride in water. The salt wash was used to remove non-specific binding of some proteins to the Protein A column that may occur; the salt wash was included in the flowthrough fraction. The elution buffer consists of 25 mM sodium phosphate and 200 mM sodium chloride at pH 2.5. The regeneration buffer consists of 25 mM sodium phosphate and 500 mM sodium chloride at pH 2.0.

### 4.2.4 SDS-PAGE Analysis of Protein A Purified Cell Lysate Fractions
Non-reducing SDS-PAGE analysis of the flow through and eluate fractions from Protein A purification of the lysates was performed to ensure that the expressed monoclonal antibodies had been effectively removed. Purified material for each of the two expressed monoclonal antibodies were included as controls. All samples were mixed with equal volumes of 2X sample buffer containing SDS and then heat denatured at 70 °C for 5 minutes in a water bath. One gel for each of the two cell lines was run. On each gel, 4 µg of the purified antibody control for each cell line, 10 µg of the flow through from each lysate and the entire eluate fraction from each lysate were loaded. The entire eluate fraction from each lysate was loaded because the amount of protein present was too low to be detected by the Bradford assay; these fractions were concentrated into as little as 5-10 µL. Electrophoresis was accomplished by subjecting the gels to a constant current of 30 mA per gel for 1 hour. The gels were then stained with coomassie blue to visualize the protein bands.

4.2.5 Deglycosylation of Glycoproteins

Prior to enzymatic treatment with N-Glycanase, all antibody-depleted fractions were buffer-exchanged into 50 mM ammonium bicarbonate in order to adjust the pH of the samples closer to the optimal pH (8.6) for N-Glycanase. One hundred microgram aliquots of each fraction were deglycosylated. The deglycosylation reaction was allowed to proceed overnight (16-20 hours) at 37 ºC. In order to isolate the released glycans, reversed phase solid phase extraction (RP-SPE) was performed using Oasis HLB 96-well SPE blocks. By performing RP-SPE, the glycans are collected in the flowthrough fraction, while the proteins are retained on the stationary phase. All glycans were evaporated to dryness in a vacuum centrifuge in preparation for derivatization.

4.2.6 Derivatization of Free Glycans with $^{12}$C$_7$ Anthranilic Acid and $^{13}$C$_7$ Anthranilic Acid
All glycan samples were derivatized by performing reductive amination. The reactions were performed with 350 mM anthranilic acid (either $^{12}$C$_7$ and $^{13}$C$_7$) in a 70:30 DMSO:acetic acid solution containing 1 M sodium cyanoborohydride. The derivatization reaction was allowed to proceed for 3 hours at 45 °C. The glycans from the first time points (100 µg each) for each of the two cell lines were derivatized with $^{12}$C$_7$ anthranilic acid, while each of the successive time points (100 µg each) were derivatized with $^{13}$C$_7$ anthranilic acid. Post-derivatization purification to remove excessive derivatization reagents was performed using Ludger Clean S cartridges, following the manufacturer’s protocol; the derivatized glycans were evaporated to dryness in a vacuum centrifuge. After drying, the glycan samples were reconstituted in purified water; and mixed with their corresponding alternately-labelled counterparts; the glycans from the first time point of each cell line were mixed with the glycans from each of the successive time points in their respective cell line. In order to further purify these glycans from potential lysate-related contaminants (e.g. hydrophilic lipids), solid phase extraction with porous graphitized carbon was performed.

4.2.7 HPLC Analysis with Fluorescence Detection of Derivatized Glycans

The derivatized glycans from the two cell lines were analyzed by anion-exchange HPLC with fluorescence detection (ex: 330 nm; em: 420 nm). The glycans were injected onto a Shodex AsahiPak NH2-P column; chromatographic separation was achieved using a two-staged linear gradient over 85 minutes. Mobile phase A is composed of 0.5 mM ammonium acetate at pH 5.0; mobile phase B is composed of 250 mM acetate at pH 5.0.

4.2.8 Quantitative MALDI-TOF MS Analysis of Isotopically-Derivatized Glycans
MALDI-TOF MS analysis was performed using an AB-Sciex 5800 TOF/TOF instrument (Foster City, CA). The glycan mixtures were spotted onto the MALDI target plate using the dried-droplet-sandwich method with MALDI-grade anthranilic acid (10 mg/mL in 50% acetonitrile-water) as the matrix. Quantitative glycomic analysis was performed by analyzing the glycan mixtures in negative reflector mode. The 1:1 ratio allows for the direct comparison and quantitation of changes in glycosylation. Each mixture was analyzed in triplicate, and final spectra were the accumulation of at least 1000 individual spectra.

4.3 Results

4.3.1 SDS-PAGE Analysis of Protein A Purified Cell Lysate Fractions

SDS-PAGE analysis was performed on the Protein A purified cell lysates to confirm the removal of the expressed antibody. This was important because the goal of the method was to investigate the glycosylation changes of all of the proteins other than the expressed biotherapeutic; the expressed antibody’s glycosylation would interfere with the glycomic analysis of the cells. For cell line A, the removal of the expressed antibody was less important as it was engineered to be aglycosyl (not glycosylated). SDS-PAGE analysis confirmed that both of the expressed antibodies had been effectively removed from their respective lysates. The flow through fractions, at each time point of both cell lines, were absent of antibody-related bands. Meanwhile, antibody-related bands were observed in some of the eluate fractions. As was previously discussed, the concentrations for the eluate fractions were not able to be determined by the Bradford assay, so it is not surprising that antibody-related bands were not observed in all of the eluate fractions. Molecular weight standards ranging from 250 down to 4 KDa were loaded into lane 1 for both gels. The purified antibody for each cell line were loaded into lane 2.
of their respective gels. The flow through fractions from the successive time points are sequentially loaded into lanes 3, 5, 7 and 9. The eluate fractions from the successive time points were sequentially loaded into lanes 2, 4, 6 and 10. In lanes 8 and 10 of the gel for cell line A, 4 bands are clearly observed, the upper-most band (~150 KDa) correspond to the intact antibody; the band below it is presumed to be a degradation product, possibly the antibody missing a light chain (HHL); the band at approximately 50 KDa is possibly free heavy chain or perhaps free Fc; the lowest clearly-visible band is possibly free light chain. The presence of free light chain can be explained by the phenomenon that antibody free heavy and light chains will tend to associate in a non-covalent manner when in solution, and therefore, free light chain could appear in the eluate fraction of a Protein A purification. For cell line B, only lanes 4 and 10 contain antibody-related bands; these bands appear to correspond to the intact antibody. The faint band directly below the antibody band is likely a degradant of some sort; possibly the clipping of a few amino acids. Refer to Figure 4.3.1-1 for details.
**Figure 4.3.1-1.** (A) Non-reduced SDS-PAGE gel for Protein A fractions of cell lysates from cell line A. The gel is loaded in the following scheme. Lane 1. Molecular weight standards. Lane 2. Control antibody for cell line A. Lane 3. Protein A flow through fraction for time point 1. Lane 4. Protein A eluate fraction for time point 1. Lane 5. Protein A flow through fraction for time point 2. Lane 6. Protein A eluate fraction for time point 2. Lane 7. Protein A flow through fraction for time point 3. Lane 8. Protein A eluate fraction for time point 3. Lane 9. Protein A flow through fraction for time point 4. Lane 10. Protein A eluate fraction for time point 4. The bands observed in lanes 8 and 10 are product-related artifacts; the upper-most band at ~150 KDa corresponds to the intact antibody; the next lower band is likely the antibody missing a light chain; the band at ~50 KDa could be free heavy chain, while the next lower band could be free light chain. (B) Non-reduced SDS-PAGE gel for Protein A fractions of cell lysates from cell line B. The gel is loaded in the following scheme. Lane 1. Molecular weight standards. Lane 2. Control antibody for cell line B. Lane 3. Protein A flow through fraction for time point 1. Lane 4. Protein A eluate fraction for time point 1. Lane 5. Protein A flow through fraction for time point 2. Lane 6. Protein A eluate fraction for time point 2. Lane 7. Protein A flow through fraction for time point 3. Lane 8. Protein A eluate fraction for time point 3. Lane 9. Protein A flow through fraction for time point 4. Lane 10. Protein A eluate fraction for time point 4. The bands observed in lanes 4 and 10 are product-related. The upper-most band at ~150 KDa is the intact antibody, while the next lower band is likely the antibody missing a few amino acids.
4.3.2 HPLC Analysis of Derivatized Glycans from Cell Lysates

By analyzing the glycans from the cell lysates by HPLC with fluorescence detection, the types of glycans present as well as the identities of some glycans can be determined. This is helpful in both providing a preview of what glycans can be expected to be observed by MALDI-TOF MS analysis and providing an orthogonal method to aid in peak identification. The chromatograms for all of the lysates from both cell lines indicated that the glycans present were non-sialylated and appeared to be mostly high mannose glycans with some neutral biantennary complex-type glycans as well. A representative chromatogram is displayed in Figure 4.3.2-1.

![Representative chromatogram of anthranilic acid-labelled glycans from cell lysates. There are five distinct regions for this particular chromatographic method. The predominant glycans observed are in the retention-time window corresponding to neutral complex-type and high mannose glycans; small peaks are observed in the region corresponding to mono-sialylated complex-type glycans. No peaks are observed in the regions corresponding to di-, tri- and tetra-sialylated complex-type glycans.](image)

**Figure 4.3.2-1.** Representative chromatogram of anthranilic acid-labelled glycans from cell lysates. There are five distinct regions for this particular chromatographic method. The predominant glycans observed are in the retention-time window corresponding to neutral complex-type and high mannose glycans; small peaks are observed in the region corresponding to mono-sialylated complex-type glycans. No peaks are observed in the regions corresponding to di-, tri- and tetra-sialylated complex-type glycans.
4.3.3 Quantitative Glycomic Analysis of CHO Cells

Following cell lysis and affinity removal of the expressed biotherapeutics, 100 µg aliquots from each of the time points were deglycosylated by treatment with N-Glycanase. The resulting glycan pools were derivatized with isotopic analogs of anthranilic acid. The glycans isolated from the first time point of each cell line were labelled with $^{12}\text{C}_7$ anthranilic acid, while the glycans isolated from the successive time points of each cell line were labelled with $^{13}\text{C}_7$ anthranilic acid. The $^{12}\text{C}_7$ anthranilic acid-labelled glycans of the first time points were mixed at a 1:1 ratio with the $^{13}\text{C}_7$ anthranilic acid-labelled glycans from each of the successive time points in their respective cell lines; this allowed any changes in glycosylation relative to the first time point to be observed. Upon the initial attempt of analysis by MALDI-TOF MS, background interference was observed in the mass spectra; the interfering peaks were of masses not believed to be related to glycans. The peaks were of consistent masses and without a corresponding 7-Da shift; this confirmed that these peaks were unrelated to the glycans. A possible explanation is that hydrophilic lipids from the cell lysates followed the glycans throughout all of the sample preparation steps. The presence of these lipids is not believed to have interfered with either the deglycosylation step or the derivatization step, as these reactions are highly specific. N-glycanase specifically releases $N$-linked glycans from glycosylated proteins and peptides; the presence of non-glycosylated species does not affect its activity. The reductive amination reaction specifically labels the reducing end of glycans with anthranilic acid; the presence of entities not capable of undergoing this reaction does not affect the reaction. In order to specifically purify these glycans from the interfering background, solid phase extraction with porous graphitized carbon (PGC) was performed, as it had previously been described as a means for specifically purifying glycans away from numerous contaminants (Packer et al 1998). The
purification with PGC was successful, as spectra from the subsequent analysis by MALDI-TOF MS were free of the interfering background.

Quantitative glycomic analysis of the CHO cells was then able to proceed. The same glycan masses were observed for both cell lines, which was not an unexpected observation as the cell lines are similar. These masses were identified as being mostly high mannose glycans with some neutral biantennary complex-type and some sialylated biantennary complex-type; glycans were identified using an in-house glycan database and cross-referenced using the glycan database available through the Consortium for Functional Glycomics (CFG; www.functionalglycomics.org). Fourteen of these glycans were selected for monitoring of changes in abundance over the course of the bioreactor campaigns; differences in glycomic changes between the two cell lines were observed. For cell line A, it was observed that the abundance of most (12) of the glycans decreased, while one glycan increased and one glycan remained unchanged at time point 2. At time point 3, all of the glycans were observed to show an increase relative to their abundances at time point 2; not all of these increases translated to an increase over the initial abundances, as seven of the fourteen glycans continued to display ratios of less than 1.0. At time point 4, most (10) of the glycans were observed to decrease in abundance relative to time point 3, while three remained relatively unchanged (≤ 5% relative increase or decrease) and one increased slightly. Overall, eight of the fourteen selected glycans show a terminal decrease in abundance relative to time point 1, four show a terminal increase and two have no terminal change. See Table 4.3.3-1 for complete details and Figure 4.3.3-1 for spectrum corresponding to the terminal change.
### CHO Cells A – Glycomic Changes During Bioreactor Campaign

<table>
<thead>
<tr>
<th>Mass (M/Z)</th>
<th>Glycan Identity</th>
<th>T2 vs. T1 Ratio</th>
<th>T3 vs. T1 Ratio</th>
<th>T4 vs. T1 Ratio</th>
<th>Terminal %Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>1014</td>
<td>Man$_2$GlcNAc$_2$Fuc</td>
<td>1.15</td>
<td>1.76</td>
<td>1.80</td>
<td>80% increase</td>
</tr>
<tr>
<td>1030</td>
<td>Man$_3$GlcNAc$_2$</td>
<td>1.00</td>
<td>1.32</td>
<td>1.30</td>
<td>30% increase</td>
</tr>
<tr>
<td>1177</td>
<td>Man$_3$GlcNAc$_2$Fuc</td>
<td>0.87</td>
<td>1.06</td>
<td>0.92</td>
<td>8% decrease</td>
</tr>
<tr>
<td>1193</td>
<td>Man$_2$GlcNAc$_2$</td>
<td>0.84</td>
<td>1.01</td>
<td>1.07</td>
<td>7% increase</td>
</tr>
<tr>
<td>1355</td>
<td>Man$_3$GlcNAc$_2$</td>
<td>0.88</td>
<td>1.14</td>
<td>1.07</td>
<td>7% increase</td>
</tr>
<tr>
<td>1380</td>
<td>Man$_3$GlcNAc$_2$Fuc</td>
<td>0.88</td>
<td>0.94</td>
<td>0.90</td>
<td>10% decrease</td>
</tr>
<tr>
<td>1517</td>
<td>Man$_3$GlcNAc$_2$</td>
<td>0.77</td>
<td>0.86</td>
<td>0.72</td>
<td>28% decrease</td>
</tr>
<tr>
<td>1583</td>
<td>G0F</td>
<td>0.82</td>
<td>1.09</td>
<td>1.01</td>
<td>No change</td>
</tr>
<tr>
<td>1679</td>
<td>Man$_3$GlcNAc$_2$</td>
<td>0.82</td>
<td>0.94</td>
<td>0.85</td>
<td>15% decrease</td>
</tr>
<tr>
<td>1786</td>
<td>Man$_3$GlcNAc$_5$Fuc</td>
<td>0.85</td>
<td>1.15</td>
<td>1.03</td>
<td>No change</td>
</tr>
<tr>
<td>1800</td>
<td>Man$_3$GlcNAc$_1$</td>
<td>0.79</td>
<td>1.18</td>
<td>0.94</td>
<td>6% decrease</td>
</tr>
<tr>
<td>1841</td>
<td>Man$_3$GlcNAc$_2$</td>
<td>0.78</td>
<td>0.87</td>
<td>0.74</td>
<td>26% decrease</td>
</tr>
<tr>
<td>1907</td>
<td>G2F</td>
<td>0.78</td>
<td>0.92</td>
<td>0.76</td>
<td>24% decrease</td>
</tr>
<tr>
<td>2003</td>
<td>Man$_3$GlcNAc$_2$</td>
<td>0.81</td>
<td>0.92</td>
<td>0.78</td>
<td>22% decrease</td>
</tr>
</tbody>
</table>

**Table 4.3.3-1.** Quantitative glycomics data for CHO cell line A. Ratios at each time point versus time point 1 for the 14 selected glycans are presented. T1=time point 1. T2=time point 2. T3=time point 3. T4=time point 4. The “Terminal %Change” corresponds to the change in abundance at time point 4 relative to time point 1. Twelve of the glycans show an initial decrease in abundance at time point 2; one shows an increase; one is unchanged. All of the glycans are observed to increase at time point 3 relative to time point 2; seven increase to levels higher than time point 1. At time point 4, ten of the glycans decrease, three are unchanged and one increases relative to time point 3. Eight of the glycans are observed to have a terminal decrease; four of the glycans are observed to have a terminal increase; two are observed to have no terminal change.
Figure 4.3.3-1. MALDI-TOF MS spectrum of glycans released from lysates of cell line A in negative reflector mode; displayed are the isotopically-labelled glycans from time point 1 and time point 4 mixed at a 1:1 ratio. The glycans from time point 1 are labelled with $^{12}$C$_7$ anthranilic acid, while the glycans from time point 4 are labelled with $^{13}$C$_7$ anthranilic acid. The fourteen selected glycans have been highlighted; the average observed ratios, displayed as $^{13}$C$_7$:$^{12}$C$_7$, are underneath the cartoon representations of the glycans. ◀:Fucose. □:N-acetylglucosamine. ●:Mannose. ○:Galactose. ◆:Sialic acid (Neu5Ac).
For cell line B, all of the glycans were observed to undergo a large initial decrease in abundance at time point 2. At time point 3, all of the glycans were observed to increase relative to time point 2; many of these increases were even more dramatic than the decreases observed at time point 2 with relative increases of up to greater than 500%; most of these increases did not result in levels above those at time point 1, as only three of the glycans increased to ratios above 1.0. At time point 4, most (9) of the glycans continued to increase in abundance, three decreased, and two remained unchanged relative to time point 3. Ultimately, seven of the glycans displayed a terminal decrease in abundance relative to time point 1, six display a terminal increase and one has no terminal change. See Table 4.3.3-2 for complete details and Figure 4.3.3-2 for spectrum corresponding to the terminal change. While the impact of these results is not currently known, they do illustrate that glycomic changes are occurring within CHO cells during a bioreactor campaign and that this method is capable of quantitating those changes.
### CHO Cells B – Glycomic Changes During Bioreactor Campaign

<table>
<thead>
<tr>
<th>Mass (M/Z)</th>
<th>Glycan Identity</th>
<th>T2 vs. T1 Ratio</th>
<th>T3 vs. T1 Ratio</th>
<th>T4 vs. T1 Ratio</th>
<th>Terminal %Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>1014</td>
<td>Man$_2$GlcNAc$_2$Fuc</td>
<td>0.35</td>
<td>1.15</td>
<td>1.46</td>
<td>46% increase</td>
</tr>
<tr>
<td>1030</td>
<td>Man$_3$GlcNAc$_2$</td>
<td>0.38</td>
<td>1.20</td>
<td>1.47</td>
<td>47% increase</td>
</tr>
<tr>
<td>1177</td>
<td>Man$_3$GlcNAc$_2$Fuc</td>
<td>0.23</td>
<td>1.18</td>
<td>1.69</td>
<td>69% increase</td>
</tr>
<tr>
<td>1193</td>
<td>Man$_4$GlcNAc$_2$</td>
<td>0.33</td>
<td>0.95</td>
<td>1.19</td>
<td>19% increase</td>
</tr>
<tr>
<td>1355</td>
<td>Man$_5$GlcNAc$_2$</td>
<td>0.21</td>
<td>0.96</td>
<td>1.46</td>
<td>46% increase</td>
</tr>
<tr>
<td>1380</td>
<td>Man$_3$GlcNAc$_2$Fuc</td>
<td>0.15</td>
<td>0.93</td>
<td>1.35</td>
<td>35% increase</td>
</tr>
<tr>
<td>1517</td>
<td>Man$_6$GlcNAc$_2$</td>
<td>0.15</td>
<td>0.55</td>
<td>0.49</td>
<td>51% decrease</td>
</tr>
<tr>
<td>1583</td>
<td>G0F</td>
<td>0.27</td>
<td>0.57</td>
<td>0.72</td>
<td>28% decrease</td>
</tr>
<tr>
<td>1679</td>
<td>Man$_7$GlcNAc$_2$</td>
<td>0.28</td>
<td>0.63</td>
<td>0.74</td>
<td>26% decrease</td>
</tr>
<tr>
<td>1786</td>
<td>Man$_3$GlcNAc$_5$Fuc</td>
<td>0.23</td>
<td>0.57</td>
<td>0.61</td>
<td>39% decrease</td>
</tr>
<tr>
<td>1800</td>
<td>Man$_9$GlcNAc$_1$</td>
<td>0.21</td>
<td>0.52</td>
<td>0.42</td>
<td>58% decrease</td>
</tr>
<tr>
<td>1841</td>
<td>Man$_9$GlcNAc$_2$</td>
<td>0.26</td>
<td>0.51</td>
<td>0.49</td>
<td>51% decrease</td>
</tr>
<tr>
<td>1907</td>
<td>G2F</td>
<td>0.41</td>
<td>0.92</td>
<td>0.95</td>
<td>No change</td>
</tr>
<tr>
<td>2003</td>
<td>Man$_9$GlcNAc$_2$</td>
<td>0.31</td>
<td>0.49</td>
<td>0.45</td>
<td>55% decrease</td>
</tr>
</tbody>
</table>

**Table 4.3.3-2.** Quantitative glycomics data for CHO cell line B. Ratios at each time point versus time point 1 for the 14 selected glycans are presented. T1=time point 1. T2=time point 2. T3=time point 3. T4=time point 4. The “Terminal %Change” corresponds to the change in abundance at time point 4 relative to time point 1. All fourteen of the glycans show an initial decrease in abundance at time point 2. At time point 3, all of the glycans were observed to show sharp increases relative to time point 2 with some showing multiple-fold increases. At time point 4, nine of the glycans continued to increase relative to time point 3, three decreased and two were unchanged. Seven of the glycans displayed a terminal decrease in abundance, six displayed a terminal increase and one showed no terminal change.
Figure 4.3.3-2. MALDI-TOF MS spectrum of glycans released from lysates of cell line B in negative reflector mode; displayed are the isotopically-labelled glycans from time point 1 and time point 4 mixed at a 1:1 ratio. The glycans from time point 1 are labelled with $^{12}$C$_7$ anthranilic acid, while the glycans from time point 4 are labelled with $^{13}$C$_7$ anthranilic acid. The fourteen selected glycans have been highlighted; the average observed ratios, displayed as $^{13}$C$_7$:$^{12}$C$_7$, are underneath of the cartoon representations of the glycans. 

- Fucose.
- N-acetylglucosamine.
- Mannose.
- Galactose.
- Sialic acid (Neu5Ac).
4.4 Discussion

A MALDI-TOF MS method for the quantitative analysis of glycomic changes occurring in CHO cells during a bioreactor campaign has been presented. The data provided allows for the monitoring of changes occurring to a single glycan, a particular subset of glycans, or [as presented] the entire glycome. This type of flexibility is beneficial as changes to a specific glycan or group of glycans may prove to be correlated to cell health or product quality of the expressed biotherapeutic. By combining this glycomic data with other information such as the proteomic data reported by Carlage and colleagues (Carlage et al 2009) and data pertaining to cell health parameters, such as cell density, viability, specific productivity, etc. to potentially draw correlations that could give some further insight into process control. It may also be possible to combine this information with product quality attributes of the expressed biotherapeutic to possibly determine if host cell protein glycosylation has any correlation to product quality.

The successful application of the method to the investigation of the CHO glycome provided support for the appropriateness of the method to be applied towards clinical glycomics and biomarker discovery. Thus the feasibility of the method to be used in such a manner will be presented in Chapter 5.

4.5 References


Crowell CK, Grampp GE, Rogers GN, Miller J, Scheinman RI. 2007. Amino acid and manganese supplementation modulates the glycosylation state of erythropoietin in a CHO culture system. *Biotechnology and Bioengineering* 96:538-49


Presta LG. 2008. Molecular engineering and design of therapeutic antibodies. *Current Opinion in Immunology* 20:460-70


Yoon SK, Choi SL, Song JY, Lee GM. 2005. Effect of culture pH on erythropoietin production by Chinese hamster ovary cells grown in suspension at 32.5 and 37.0 ºC. *Biotechnology and Bioengineering* 89:345-56
Chapter 5

A GENERAL APPROACH FOR THE PURIFICATION AND QUANTITATIVE GLYCOMIC ANALYSIS OF HUMAN PLASMA

Contributing Authors: Samnang Tep, Marina Hincapie and William S. Hancock

Submitted for publication in Analytical and Bioanalytical Chemistry

Human plasma provided through NIH Grant U01-CA128427. A sincere thank you to the NIH.
Chapter 5

A GENERAL APPROACH FOR THE PURIFICATION AND QUANTITATIVE GLYCOMIC ANALYSIS OF HUMAN PLASMA

Submitted for publication in Analytical and Bioanalytical Chemistry

5.1 Introduction

Over time it has been realized that glycosylation is closely connected to various diseases. Some diseases are directly attributed to deficiencies in glycosylation or glycosylation pathways; these diseases have been termed Congenital Disorders of Glycosylation (CDG) (Aebi et al 1999). Two groups, type I and type II, have been established with lower case letters distinguishing the subgroups (e.g. type Ia, type IIb, etc.) correlating to the chronological identification of the defective gene. All of these diseases are autosomal recessive and hallmarked by a general inability to thrive (Ohtsubo & Marth 2006); the mortality rate is 20% within the first five years of life (Freeze 2006). The first case of CDG was described in 1980 (Jaeken & Matthijs 2007); since then a total of forty-five CDG have been identified (Jaeken 2010); nineteen of which are disorders affecting N-glycosylation (Freeze 2006). Twelve of these nineteen are type I and seven are type II. The twelve type I CDG all display hypo- or underglycosylated proteins, which has been linked to either the insufficient supply or incomplete synthesis of the Dol-P-P-GlcNAc$_2$Man$_9$Glc$_3$ glycan. The result of this is that all twelve type I CDG express proteins with unoccupied glycosylation sites and/or incorrect glycosylation leading to loss(es) of function. Type Ia is the most common CDG with more than 700 known patients (Jaeken 2010). Mutations (90 mutations have been identified) to the $PMM2$ gene that codes for one of the two phosphomannosmutase enzymes responsible for the conversion of mannose-6-phosphate to
mannose-1-phosphate have been identified as the underlying cause. Because mannose-1-phosphate is the precursor for GDP-mannose, which donates mannose residues for attachment to the growing Dol-P-P-GlcNAc₂Man₃ glycan, its disruption affects the glycosynthetic pathway at a very early stage. Common characteristics include inverted nipples, unusual subcutaneous fat deposits, axial hypotonia, internal strabismus (crossed-eyes) and severe psychomotor impairment (Marquardt & Denecke 2003). Other type I CDG are far less common than type Ia; there are fewer than 100 known patients in total. Type II CDG are distinguished from Type I as they are caused by defects or alterations in the processing of the glycans after they have been successfully attached to proteins and transferred to the golgi apparatus (Freeze 2006). The resulting symptoms for type II CDG-affected patients vary widely and very few patients are known to be affected by these diseases. Only one patient, each, has been identified as having type IIb and type IIId; the patient affected with type IIb died at 2.5 months.

Other diseases with observed changes in glycosylation, not considered CDG, are often regarded as acquired diseases. These include various cancers (Dall'olio 1996), liver diseases (Blomme et al 2009; Comunale et al 2010), Alzheimer’s disease (Botella-Lopez 2006; Charlwood 2001; Liu et al 2002), schizophrenia (Bauer et al 2010; Stanta et al 2010), muscular dystrophies (Manya et al 2009; Martin-Rendon & Blake 2003), diabetes (Carlsson et al 2008; Itoh et al 2007) and many others. The changes in glycosylation in these diseases varies, allowing for some discernable differences between diseases to be observed and the potential to identify glycosylation-related biomarkers. For example, many cancers have been observed to have increased levels of branching and sialylation (Dall'olio 1996); colon cancer has been repeatedly reported to present with increased sialylation, specifically α-2,6 linked (Dall'Olio et al 1991; Dall'Olio et al 1992a; b; Saitoh et al 1992). In one report, it was presented that approximately 90% of the samples of
colon carcinomas tested displayed increased α-2,6 linked sialylation (Dall'Olio et al 1989). Breast (Kyselova et al 2007), ovarian (Saldova et al 2008) and cervical cancers (López-Morales et al 2010) have all been reported to display increased sialylation as well. Meanwhile, rheumatoid arthritis is typically characterized by a decrease in IgG galactosylation (Kratz et al 2009; Matsumoto et al 2000); this has been observed to the extent that it has been recognized as a marker, not only for the presence of rheumatoid arthritis, but also for the progression of the disease (Arnold et al 2007).

In recent years, some efforts have been made to develop methods to quantitate these changes in glycosylation in order to provide information that would possibly lead to the identification of biomarkers. These methods range from employing isotopic labelling in vivo (Orlando et al 2009) to isotopic permethylation of released glycans (Alvarez-Manilla et al 2007; Atwood et al 2008) to the use of in-house chemically synthesized discernable mass tags (Bowman & Zaia 2007). While these methods have helped to make large strides in the field of glycomics, they can be quite involved, require highly sensitive mass spectrometers and generate very complex data. In order to perform in vivo labelling (Orlando et al 2009), the cells must be cultured in differential media containing either $^{14}\text{N}$ and $^{15}\text{N}$ isotopes of glutamine. The incorporation of $^{15}\text{N}$-glutamine into the resulting glycans generates a series of mass differences depending on the size and type of the glycan (number of GlcNAc units); high mannose glycans would be separated by only 2 Da, while complex-type glycans would be separated by as few as 4 Da (biantennary) up to 6 Da (tetra-antennary) or greater (lactosamine repeats). This series of mass differences results in greater complexity of the mass spectra and a higher potential for overlapping masses; further, the smaller mass differences of 2-4 Da do not provide a large enough mass difference for quantitative analysis by lower-resolving mass spectrometers. In the methods utilizing isotopic
permethylation (Alvarez-Manilla et al 2007; Atwood et al 2008), permethylation of glycans is performed by the method described by Ciucanu and Kerek in 1984 (Ciucanu & Kerek 1984). The method requires treatment of the glycans with a very strong base and performing organic extractions to isolate the permethylated glycans. The permethylation of glycans can also result in the loss important modifications, such as phosphorylation or sulfation. Because permethylation modifies all of the available hydroxyl groups on the glycans, the mass differences are not consistent between all glycan pairs; this creates more complexity in the mass spectra and can potentially lead to overlapping masses. In the method developed by Bowman and Zaia (Bowman & Zaia 2007), solid phase synthesis was used to generate discernable mass tags; this type of reagent preparation can be cumbersome and requires some expertise in organic synthesis. The mass tags generated consisted of a series of four tags separated by 4 Da, creating a normal, a +4, a +8 and a +12 mass difference. The 4 Da mass difference between the tags may lead to overlapping masses between samples. Ideally, the method should incorporate minimal modifications to the glycans, employ well-established chemistries, utilize commercially available reagents and deliver a wide dynamic range with high sensitivity. Additionally, the implementation of $^{13}$C rather than $^2$H is preferred in order to preserve the chromatographic properties of the derivatized glycans (Julka & Regnier 2004). A method fitting these criteria was previously developed (Xia et al 2009); the method utilized $^{12}$C$_6$ and $^{13}$C$_6$ analogs of aniline to derivatize released glycans. The authors reported good data accuracy, reproducibility and linearity; the authors also presented that the derivatized glycans were amenable to analysis by either LC-ESI MS or MALDI-TOF MS. The drawbacks with the method are the use of a UV absorbing tag rather than a fluorescence tag and the typical requirement of performing MALDI-TOF MS analysis in two different modes (positive and negative) with two different matrices to
obtain data on neutral and sialylated glycans. The method presented in Chapter 2, and employed here, is advantageous over the method presented by Xia and colleagues. Those advantages were discussed in Chapter 2. Briefly, the use of the highly fluorescent tag, anthranilic acid, to derivatize the glycans instead of the UV-absorbing aniline allows for fluorescence detection, which can be up to 1000 times greater than UV detection; Additionally, the use of anthranilic acid as the tag and matrix allows for the simultaneous analysis of neutral and sialylated glycans, which streamlines analysis by removing the requirement of using a different matrix and acquisition mode for neutral and sialylated glycans with MALDI-TOF MS analysis.

Also discussed here are a series of steps that is believed to be an effective methodology that can be generally applied to prepare human plasma samples for quantitative glycomic analysis. The first step is the depletion of plasma of the highest abundant proteins; depletion of highest abundant proteins is important, as it is generally agreed that in order to investigate biofluids for biomarkers, the complexity of the material must be reduced (Ahmed 2009). This is accomplished using the Agilent Multiple Affinity Removal System (MARS) to remove the 14 highest abundant proteins from plasma. The second step is the enrichment of the depleted plasma for glycoproteins via Multi-Lectin Affinity Chromatography (MLAC) (Yang & Hancock 2004; 2005); in this methodology the selected lectins have been modified to target a more specific glycosylation.

5.2 Experimental

5.2.1 Reagents

The 14 protein multi-affinity removal (14P MARS) column, the corresponding mobile phases and Poros AL powder were purchased from Agilent Technologies (Santa Clara, CA). Isolated
lectins (were purchased from Vector Labs (Burlingame, CA). Human plasma, normal and diseased (stage 4 lung cancer), were acquired through NIH grant U01-CA128427. N-Glycanase (PNGase F; EC 3.5.1.52) was purchased from Prozyme (Hayward, CA). Anthranilic acid, dimethyl sulfoxide and sodium cyanoborohydride were purchased from Sigma Aldrich (St. Louis, MO). The $^{13}$C$_7$ isotopic analog of anthranilic acid was purchased from IsoSciences (King of Prussia, PA). Acetic acid, acetonitrile and HPLC-grade water were purchased from Fisher Scientific (Pittsburg, PA). Post-derivatization purification cartridges (LudgerClean S) were purchased from Ludger, Ltd. (Oxfordshire, UK). Solid phase extraction blocks (Oasis HLB) were purchased from Waters Corporation (Milford, MA). Centrifugal filtration devices (5,000 molecular weight cut off) were purchased from Millipore (Billerica, MA).

5.2.2 Depletion of Human Plasma by MARS

In order to remove 14 of the highest abundant proteins, the human plasma samples were separated over an Agilent MARS 14P column. The MARS 14P column has specific affinities for 14 of the proteins in highest abundance in human plasma and sera. Both level 1 and level 2 proteins are depleted. Level 1 proteins depleted include albumin, transferrin, haptoglobin, IgG, IgA, and $\alpha$1-antitrypsin; level 2 proteins depleted include fibrinogen, $\alpha$2-macroglobulin, $\alpha$1-acid glycoprotein, complement C3, IgM, apolipoprotein AI, apolipoprotein AII, and transthyretin. Performing this depletion removes 94% of all proteins present in serum and plasma, substantially reducing the complexity and background interference created by the higher abundant proteins. The method uses a two-buffer system; the buffer compositions are proprietary formulations developed by Agilent Technologies. Buffer A is a neutral salt-containing buffer, while Buffer B is a low pH urea-containing buffer that elutes all of the affinity-bound proteins from the column.
Just prior to loading onto the column, the plasma samples were diluted by a factor of four with Buffer A and then passed over a 0.22 µm filter to remove any particulate matter. In order to not exceed the column’s capacity (20 µL serum or plasma for 4.6 mm x 50 mm column), individual 80 µL injections of the diluted plasma samples were made until all of the material had been depleted. The flowthrough fractions, containing the lower abundant proteins, were collected for each run and pooled. The pooled fractions were buffer exchanged (MLAC buffer A) and concentrated by centrifugation over 5,000 molecular-weight cut off filter devices. The concentrations of the pooled/concentrated flowthrough fractions were determined using the Bradford assay.

5.2.3 Enrichment of Glycoproteins by MLAC

In order to enrich for glycoproteins in the 14P-depleted plasma samples, the samples were separated over an MLAC column. The column used was a modified version of the most recent iteration of the MLAC technology, HP-MLAC (Kullolli et al 2010), in that the selected lectins were different; the lectin conjugation and column packing procedures were maintained. The selected lectins were modified to more specifically target differences in sialylation. The original MLAC lectins included concanavalin A (ConA; *Canavalia ensiformis*), wheat germ agglutinin (WGA; *Triticum vulgaris*) and jacalin (Jacalin; *Artocarpus integrifolia*). ConA has specificity towards alpha-linked mannoses; WGA has specificity towards N-acetylglucosamines; Jacalin has specificity towards O-linked glycans containing galactoses β-1,3 linked to N-acetylglucosamine. The lectins included in the modified MLAC column are ConA, *Erythrina cristagalli* lectin (ECL; *Erythrina cristagalli*), *Sambucus nigra* lectin (SNA; *Sambucus nigra*) and *Maackia amurensis* lectin II (MAL-II; *Maackia amurensis*). ECL has specificity towards terminal galactoses β-1,4 linked to N-acetylglucosamine; SNA has specificity towards terminal sialic
acids α-2,6 linked to galactose; MAL-II has specificity towards terminal sialic acids α-2,3 linked to galactose. The incorporation of two sialic acid binding lectins is hypothesized to better capture sialylated glycans, whether normally sialylated or hypersialylated due to disease. The inclusion of a terminal galactose binding lectin, ECL, is to capture both normally non-sialylated glycans as well as those glycans that may have become desialylated due to disease. The method uses a three-buffer system. Buffer A contains 25 mM Tris, 0.5 M NaCl, 1 mM MnCl₂, 1 mM CaCl₂ and 0.05% sodium azide at pH 7.4; buffer B contains 100 mM acetic acid to elute the bound glycoproteins from the column; buffer C contains 500 mM Tris, 1 M NaCl, 0.05% sodium azide at pH 7.2 to regenerate the column.

The depleted plasma samples were loaded onto the modified MLAC column and separated using a step gradient. The eluted fractions, containing the enriched glycoproteins, were immediately neutralized using 1 M ammonium bicarbonate. The eluates were pooled and then buffer exchanged (50 mM ammonium bicarbonate) and concentrated by centrifugation over 5,000 molecular-weight cut off filter devices. The concentrations of the pooled/concentrated flowthrough fractions were determined using the Bradford assay.

5.2.4 Deglycosylation of Glycoproteins

The enriched glycoproteins (50 µg each) from plasma were treated with N-Glycanase at 37 °C for 16 to 20 hours (overnight) to remove the N-linked glycans. Following deglycosylation, the released glycans were isolated via reversed phase solid phase extraction using Oasis HLB 96-well blocks. By performing reversed phase solid phase extraction, the released glycans are collected in the flowthrough, while the deglycosylated protein is bound to the stationary phase. This procedure is straightforward as it avoids using a bind/elute protocol that may result in the
loss of some glycans if they are not completely eluted from the stationary phase. The flowthroughs were evaporated to dryness in a vacuum centrifuge in preparation for derivatization.

5.2.5 Derivatization of Free Glycans with $^{12}$C$_7$ Anthranilic Acid and $^{13}$C$_7$ Anthranilic Acid

Free glycans were derivatized by reductive amination with 350 mM anthranilic acid ($^{12}$C$_7$ or $^{13}$C$_7$), 1 M sodium cyanoborohydride in a 70:30 DMSO:acetic acid solution. The derivatization was allowed to proceed for 3 hours at 45 °C. The released glycans from 50 µg of 14P depleted/MLAC enriched normal plasma were derivatized with $^{12}$C$_7$ anthranilic acid, while the released glycans from 50 µg of 14P depleted/MLAC enriched diseased plasma were derivatized with $^{13}$C$_7$ anthranilic acid. Post-derivatization cleanup of all labeled glycans was performed using Ludger Clean S cartridges following the manufacturer’s suggested protocol. The purified, labeled glycans were evaporated to dryness in a vacuum centrifuge.

5.2.6 HPLC Analysis with Fluorescence Detection of Derivatized Glycans

The derivatized glycans from the normal and diseased plasma were analyzed by HPLC with fluorescence detection. Increased sensitivity with fluorescence (ex: 330 nm; em: 420 nm) is afforded with the use of anthranilic acid; thereby, allowing for detection of lower-abundant glycans and/or the use of less analyte. The glycans are separated over a Shodex AsahiPak NH2-P column using an increasing ammonium acetate gradient. Mobile phase A is 0.5 mM ammonium acetate at pH 5.0; mobile phase B is 250 mM ammonium acetate at pH 5.0; chromatographic separation is achieved using a two-staged linear gradient over 85 minutes. Performing this type of HPLC analysis allows for a previewing of the types of glycans present in the sample. With an understanding of where certain glycans and types of glycans elute using
this particular chromatographic method, a powerful orthogonal method can be utilized to aid in peak identification.

5.2.7 Quantitative MALDI-TOF MS of Isotopically Derivatized Glycans

MALDI-TOF MS of the derivatized glycans was performed using an AB-Sciex 5800 TOF-TOF instrument (Foster City, CA). The evaporated glycan samples were reconstituted in purified water and then mixed; the differentially labelled glycans of normal and diseased plasma were mixed at a 1:1 ratio, so that any differences in the abundance of any glycan could be observed and directly quantitated. The glycan mixtures were applied to the MALDI target plate using the dried-droplet-sandwich method; triplicate analysis was performed at each point. The matrix utilized was MALDI-grade anthranilic acid (10 mg/mL) in 50% acetonitrile-water.

5.3 Results

5.3.1 Premise and Methodology

The goal of this methodology is to provide both a general means for the preparation of glycomic samples from human plasma/sera that are simplified and enriched as well as an analytical method for quantitative analysis. The removal of 14 high abundant proteins by MARS 14P depletion followed by the specific enrichment of glycoproteins with MLAC, provides samples that are well-suited for quantitative glycomic analysis. The derivatizing of free glycans with either a normal \(^{12}\text{C}_7\) or heavy \(^{13}\text{C}_7\) tag allows for the relative quantitation between differentially labelled glycan samples. With this, changes to even a single glycan can be observed and quantitated as the data are presented as glycan pairs separated by 7 daltons; changes to a particular class or group of glycans can also be observed. For example, overall sialylation
changes can be determined by calculating the change in ratios for all sialylated glycan pairs between samples. This type of analysis can be very powerful, as it allows for the monitoring of glycosylation changes over time. The ability to monitor the change in glycosylation over the progression of disease would be immensely informative, perhaps becoming predictive or diagnostic with continued use. A workflow of the methodology is described in Figure 5.3.1-1.
Figure 5.3.1-1. Workflow for purification, preparation and analysis of samples for quantitative glycomic analysis of human plasma by MALDI-TOF MS. The first step involves removal of high abundant proteins by MARS 14P depletion. The second step is the enrichment of glycoproteins in the depleted plasma by MLAC. The third step is the deglycosylation of the enriched glycoproteins by treatment with N-Glycanase and the subsequent isolation of the released glycans. The fourth step is the derivatization of the glycans with isotopic analogs of anthranilic acid and their subsequent purification. The final step is quantitative glycomic analysis by MALDI-TOF MS.
5.3.2 High Abundant Protein Depletion Human Plasma with MARS 14P

Human plasma samples were depleted of 14 of the highest abundant proteins by purification over a muti-affinitity column (MARS 14P). Depletion of these high abundant proteins reduced the complexity of the human plasma samples by 94%, which is an important step as the reduction in background interference created by high abundant proteins is necessary for biomarker discovery (Ahmed 2009). The diseased human plasma was the result of a pooling of plasma from five patients of similar sex, age and ethnicity (male, 50-55 years, Caucasian), all afflicted with stage-4 lung cancer; the normal human plasma was the result of a pooling of plasma from five healthy volunteers of the same description. The depletion of plasma was performed according to the manufacturer’s protocol. The results indicated that the depletions were successful, as the chromatographic profiles for both normal and diseased plasma were visually comparable to the chromatographic profile provided by the manufacturer. See Figure 5.3.2-1 and -2. The flowthrough fractions were collected for MLAC enrichment.

![Figure 5.3.2-1. Example chromatogram of MARS 14P depletion of human plasma provided by the manufacturer. Figure adapted from package insert for MARS 14P column.](image-url)
Figure 5.3.2-2. Representative chromatograms for the depletion by MARS 14P of normal and diseased human plasmas. The flowthrough fractions were collected for further processing and analysis.
5.3.3 Enrichment of Glycoproteins by MLAC

Glycoprotein enrichment was achieved using a modified MLAC column. The modified column contained four immobilized lectins: ConA, ECL, SNA and MAL-II. These lectins were selected to specifically target changes in sialylation. The depleted plasma samples (normal and diseased) were separated over the modified MLAC column; chromatographic profiles were similar to previously observed profiles with standard MLAC. See Figures 5.3.3-1 and -2. The eluates were collected for further processing in preparation for quantitative glycomic analysis.

Figure 5.3.3-1. Representative chromatogram of MLAC enrichment of human plasma using the standard MLAC column. Figure adapted from (Kullolli et al 2010).
Figure 5.3.3-2. Overlay of representative chromatograms from the MLAC enrichment step for the 14P depleted flowthrough fractions. The upper trace is a blank injection (buffer A); there is a very slight peak at approximately 16 minutes, which is likely an artifact due to a change in buffer composition. The middle trace is that of the diseased plasma; the lower trace is that of normal (control) plasma. The eluate fractions in both normal and diseased plasma constituted approximately 27% of the total peak area. The eluate fractions from MLAC contain the specifically-enriched glycoproteins; these fractions were collected for further processing.
5.3.4 HPLC Analysis of Derivatized Glycans

Chromatographic analysis of the derivatized glycans was performed in order to assess what glycans or types of glycans were present in the 14P-depleted/MLAC-enriched fractions. The glycans from both the normal and diseased plasma were analyzed. The results showed fairly similar profiles between the normal and diseased samples with both containing almost exclusively sialylated glycans; neutral glycans were observed in very low abundance. This was expected as the modified MLAC column was designed to capture sialylated glycans. The predominant species observed in both sample was a disialylated glycan; based upon the retention time (~46 minutes) the identity is presumed to be the A2F glycan. The monosialylated glycan at approximately 27 minutes is presumed to be A1F, and the trisialylated glycan at approximately 63 minutes is presumed to be A3F. See Figure 5.3.4-1. This information allowed for an understanding of what types of glycans would be observed by MALDI-TOF MS analysis.
Figure 5.3.4-1. Overlay of chromatograms of derivatized glycans isolated from 14P-depleted/MLAC-enriched normal and diseased plasma samples. The top trace is that of glycans from normal plasma. The bottom trace is that of glycans from diseased plasma. Glycans of various sialylation levels dominate the profile. Putative assignments based on retention time have been made regarding the identities of one of each of the sialylated glycans. The A1F, A2F and A3F glycans have been labelled.
5.3.5 Quantitative Glycomic Analysis of Normal vs. Diseased Plasma

Following high abundant protein depletion by MARS 14P and glycoprotein enrichment by MLAC, a 50 µg aliquot of each, normal and diseased plasma, was deglycosylated by treatment with N-glycanase. The released glycans were alternately labelled with $^{12}$C$_7$ anthranilic acid (normal plasma) and $^{13}$C$_7$ anthranilic acid (diseased plasma). Following post-derivatization purification and evaporation, the differentially labelled glycans were reconstituted in purified water and mixed at a 1:1 ratio. The mixed glycans were spotted onto the MALDI plate using the dried-droplet-sandwich method and then analyzed. Numerous glycans and glycan pairs were observed; the sialylated glycans were of particular interest as it was anticipated that they would show obvious differences in abundance due to the disease state being stage-4 lung cancer. Six of the glycan pairs, which were of recognized masses, corresponded to mono-, di- and trisialylated glycans. The glycan pair at m/z 2052.09/2059.11 corresponds to the monosialylated biantennary glycan without a core fucose (A1), while the glycan pair at m/z 2199.16/2206.18 correspond to its core-fucosylated counterpart (A1F). The glycan pairs at m/z 2344.23/2351.25 and m/z 2490.28/2497.30 correspond to disialylated biantennary glycan without (A2) and with core fucosylation (A2F), respectively. The glycan pairs at m/z 3000.50/3007.53 and m/z 3146.57/3153.60 correspond to the trisialylated triantennary glycan without (A3) and with a core fucose (A3F), respectively. As expected, the ratios for all six of these sialylated glycans pairs were observed to be elevated towards a higher abundance of sialylated glycans in the diseased state. The ratios ranged from a 200% up to a 500% increase in abundance; the observed CVs for the glycan pairs were less than 10%. See Figure 5.3.5-1. These results illustrate the ability of this methodology to provide meaningful quantitative glycomic data and its appropriateness for the use in glycomic studies.
Figure 5.3.5-1. Quantitative MALDI-TOF MS analysis of normal vs. diseased plasma. Six glycan pairs, all sialylated, were focused on. An increase in the abundances of all of these glycans were observed. The ratios are presented as $^{13}$C$_7$:$^{12}$C$_7$ so that changes relative to the normal plasma can more easily be observed; values of greater than 1.0 correspond to an increase in abundance, while values less than 1.0 correspond to a decrease in abundance. The inset table displays the ratio values for the six glycan pairs. Cartoon representations of the glycan structures have been included. 

- Fucose
- N-acetylglucosamine
- Mannose
- Galactose
- Sialic acid (Neu5Ac)

<table>
<thead>
<tr>
<th>Glycan</th>
<th>Average</th>
<th>RSD</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>2.78</td>
<td>2.8%</td>
</tr>
<tr>
<td>A1F</td>
<td>2.01</td>
<td>1.2%</td>
</tr>
<tr>
<td>A2</td>
<td>3.61</td>
<td>9.1%</td>
</tr>
<tr>
<td>A2F</td>
<td>2.68</td>
<td>8.2%</td>
</tr>
<tr>
<td>A3</td>
<td>3.26</td>
<td>1.9%</td>
</tr>
<tr>
<td>A3F</td>
<td>5.08</td>
<td>2.9%</td>
</tr>
</tbody>
</table>
5.4 Discussion

A general approach for the preparation and quantitative glycomic analysis of human plasma has been developed. The methodology incorporates the depletion of high abundance plasma proteins by multi-affinity removal (MARS 14P), the enrichment of glycoproteins by multi-lectin affinity chromatography (MLAC), the isotopic labelling of glycans by reductive amination and the mass spectrometric analysis by MALDI-TOF MS. This methodology provides samples for glycomic analysis that are both less complex and enriched, which is important in the investigation and discovery of potential biomarkers.

In this particular workflow, the lectins used in the MLAC column were modified to incorporate two terminal sialic acid specific lectins and one terminal galactose specific lectin. The hypothesis behind this modification was to specifically capture glycoproteins that have undergone changes (increase or decrease) in terminal sialylation of their glycans. Examples of this type of change in terminal sialylation can be observed in cervical cancer (López-Morales et al 2010) and alcoholic liver disease (Blomme et al 2009), wherein there is an observed increase in terminal sialylation in patients with cervical cancer and an observed decrease in terminal sialylation in patients with alcoholic liver disease. Because of the sample’s disease state, stage 4 lung cancer, the modified MLAC column was anticipated to work well; the chromatographic data of the derivatized glycans and the MALDI-TOF MS data generated support this. Further, this suggests that the MLAC technology can be modified to suit more targeted glycomic analysis. For example, if a disease state is either believed to be or has been observed to be associated with increased fucosylation, a fucose specific lectin such as Aleuria aurantia lectin (AAL; Aleuria aurantia) could be incorporated into the MLAC column to specifically capture fucosylated glycans. The apparent versatility of the MLAC technology may allow for the
development of more diagnostic methods, as the selection of lectins used in the column can take advantage of established changes in glycosylation associated with certain diseases.

The type of data provided by the MALDI-TOF MS method, the relative ratio of glycan pairs, allows for the determination of glycosylation changes to a single glycan, a group of glycans, or the entire glycome. This kind of flexibility is important as the monitoring of changes to a particular glycan or type(s) of glycan(s) may yield diagnostic or prognostic information. For example, many cancers are known to show an increase in sialylation of serum glycoproteins (Reis et al 2010), while patients affected by rheumatoid arthritis have shown a decrease in terminal end-capping of neutral biantennary glycans with galactose of their IgGs (Kratz et al 2009; Matsumoto et al 2000). The use of anthranilic acid as the derivatization reagent allows for detection by fluorescence if analyzed by chromatography or capillary electrophoresis, and in so doing a preview of the glycans that are present as well as the overall relative quantitation of glycans can be achieved with minimal sample consumption; this preserves the bulk of the sample material for MALDI-TOF MS analysis.

5.5 References


Dall'Olio F, Malagolini N, Serafini-Cessi F. 1992b. The expression of soluble and cell-bound alpha 2,6 sialyltransferase in human colonic carcinoma CaCo-2 cells correlates with the degree of enterocytic differentiation. *Biochem Biophys Res Commun* 184:1405-10


Chapter 6

SUMMARY OF RESULTS AND PROSPECTIVE FUTURE APPLICATIONS

6.1 Summary of Results

In Chapters 3, 4 and 5 experiments were performed in which changes in glycosylation were observed and quantitated. There are numerous potential explanations and various implications for these changes in glycosylation.

The changes in glycosylation observed on the CHO-expressed Fc-fusion protein in Chapter three could be due to either a lower specific productivity, a decrease in cell density and viability, an increase in aggregation or misfoldings or an increase in the expression of underglycosylated and undersialylated protein. All of these potential causes could be the result of the extended time of cell culture, as cells are typically grown in a very high stress environment to encourage higher expression of the target biotherapeutic protein. The potential impact of the change in glycosylation ranges from a lower yield of the active construct or usable product, a more complicated purification process (if aggregates and misfoldings are increased), a decrease in stability of glyco-occupancy is decreased, or a change in pharmacokinetic properties related to sialylation (decrease in sialylation was observed). This type of information could be used to make modifications to cell culture. For instance, if the sialylation of the product is known to affect its pharmacokinetic properties, the data suggests that the harvest date be changed to time point 5.

The changes in glycosylation observed on the host-cell proteins in CHO cells in Chapter four could be due to changing levels of expression of either the glycosyltransferase enzymes or the
host-cell proteins themselves at various times. The glycosylation or change in glycosylation of glycan substrates by the expressed biotherapeutic protein could also potentially affect the glycosylation of the host-cell proteins. As previously discussed in Chapter 4, the impact of these changes in glycosylation are unknown, but it is reasonable to speculate that changes in the glycosylation of host-cell proteins could affect the health of the CHO cells, as studies in developmental biology have confirmed that changes in protein glycosylation occur throughout development and ageing.

The changes in glycosylation observed in diseased (lung cancer) human plasma in Chapter 5 were not unexpected, as increases in sialylation have been observed on various cancers in previous studies by others. A potential cause for this increase in sialylation could be an increase in the expression of normally-sialylated proteins such that their abundance in the blood of a diseased patient far exceeds that of a normal patient; combining this type of data with proteomic data could help to confirm or refute this hypothesis. Another potential cause could the increased expression of sialyltransferase enzymes, thereby potentially increasing the sialylation of all expressed glycoproteins. The impact of this data is that another type of cancer is observed to have an increase in sialylation, and that a method capable of quantitating that change has been developed.

6.2 Prospective Future Applications

The goal of developing a method capable of providing meaningful quantitative data regarding changes in glycosylation for industrial or academic and clinical purposes has been accomplished. The method utilized well-established chemistries with commercially-available reagents and commonly-used instrumentation. Its universal applicability was displayed by the successful
analysis of various types of samples with varying levels of complexity. The inherent flexibility of the method allows for the quantitation of a single glycan, a group of glycans or the entire complement of glycans present within a complex biological sample. The method was presented as stand-alone method, but the use of isotopic-carbon analogs of anthranilic acid to derivatize the glycans makes them suitable for hyphenated methods. Because the isotopic analogs do not chromatographically resolve, they can be simultaneously eluted from a column and introduced into a mass spectrometer by electrospray ionization (LC-MS) or automatically spotted onto a MALDI target plate (LC-MALDI). Further, tandem mass spectrometry (MS/MS) can be performed on the glycans for structural/linkage investigations.

Because of its generality, there are a number of prospective applications for this method. Having displayed its ability to be used as an analytical tool for the quantitation of glycosylation changes in a disease state, the method could be used for various glycomic investigations and biomarker discovery. Additionally, the method could be applied to developmental biology studies in order to monitor and quantitate the glycomic changes occurring during biological development. Within the biopharmaceutical industry, the method could be used to monitor the glycosylation changes on a product as it relates to the various process steps and their conditions, ranging from cell line screening and selection, to cell culture development, to product purification, to formulation development and even to product stability. Additionally, host cell glycomics could be investigated to determine whether host cell protein glycosylation has any correlation to cell health and/or product quality. In fact, the implementation of these types of applications is just on the horizon at Biogen Idec.

As with most any method, there is room for improvement. A couple of points of improvement can be found with sample preparation; some of the preparative steps are fairly time consuming.
The deglycosylation step is an overnight (16-24 hours) digest. There are a number of possible ways to shorten the procedure. A product used at Biogen Idec, Rapid Deglycosylation Kits (Prozyme, Hayward, CA), reduces deglycosylation time down to 1 hour. The limitation of the kits is that they have only been qualified for use on singular purified substances; also, the kit has not worked well with some proteins. Another option would be to perform the deglycosylation step following a proteolysis as the deglycosylation of peptides is faster than the deglycosylation of proteins. However, proteolysis steps often require lengthy incubations as well; the combined time for proteolysis and deglycosylation could end up being very similar to 16-24 hours. The glycan derivatization step is 3 hours (at 45 ºC). This could be expedited by increasing the incubation temperature, but doing so could result in the desialylation of sialylated glycans. Another option would be use a different labelling chemistry; a product investigated by Biogen Idec, Instant-AA (Prozyme, Hayward, CA), derivatized glycans in the glycosylamine form and accomplished labelling within seconds. The product is limited by not being available in an isotopic format; the actual identity of the tag and chemistry are not disclosed by the manufacturer. An additional point for improvement would be to mix the isotopically-labelled glycans following derivatization prior to performing the post-derivatization purification, as this should increase data accuracy by removing a potential variable.