

University of New Hampshire  
University of New Hampshire Scholars' Repository

---

Master's Theses and Capstones

Student Scholarship

---

Fall 2009

# N-linked glycans characterization and biomarker discovery by sequential mass spectrometry

Hongqin (Jenny) Jiao

*University of New Hampshire, Durham*

Follow this and additional works at: <https://scholars.unh.edu/thesis>

---

## Recommended Citation

Jiao, Hongqin (Jenny), "N-linked glycans characterization and biomarker discovery by sequential mass spectrometry" (2009). *Master's Theses and Capstones*. 483.

<https://scholars.unh.edu/thesis/483>

This Thesis is brought to you for free and open access by the Student Scholarship at University of New Hampshire Scholars' Repository. It has been accepted for inclusion in Master's Theses and Capstones by an authorized administrator of University of New Hampshire Scholars' Repository. For more information, please contact [nicole.hentz@unh.edu](mailto:nicole.hentz@unh.edu).

**N-LINKED GLYCANS CHARACTERIZATION AND  
BIOMARKER DISCOVERY BY SEQUENTIAL MASS  
SPECTROMETRY**

**BY**

**HONGQIN (JENNY) JIAO**

**B.S., Xi'an Jiaotong University, P.R. China, 1994**

**THESIS**

**Submitted to the University of New Hampshire  
in Partial Fulfillment of  
the Requirements for the Degree of**

**Master of Science  
in  
Chemistry**

**September, 2009**

UMI Number: 1472068

### INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

**UMI<sup>®</sup>**

---

UMI Microform 1472068  
Copyright 2009 by ProQuest LLC  
All rights reserved. This microform edition is protected against  
unauthorized copying under Title 17, United States Code.

---

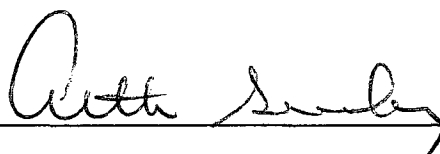
ProQuest LLC  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106-1346

This thesis has been examined and approved.



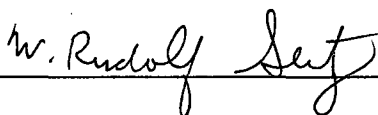
---

Thesis Director, Vernon N. Reinhold,  
Professor of Biochemistry and Molecular Biology



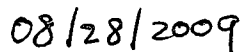
---

Arthur Greenberg, Professor of Chemistry



---

W. Rudolf Seitz, Professor of Chemistry



---

Date

# TABLE OF CONTENTS

LIST OF TABLES.....	v
LIST OF FIGURES.....	vi
ABSTRACT.....	viii

CHAPTER	PAGE
1. INTRODUCTION	
1.1 Glycans and Glycoproteins.....	1
1.2 N-linked Glycans.....	2
1.3 Analytical Tools for Glycan Structures.....	5
2. ISOMERS CHARACTERIZATION OF N-LINKED GLYCANS FROM CHICKEN OVALBUMIN BY MS <sup>n</sup> .....	9
2.1 Introduction.....	9
2.2 Experimental Methods .....	10
2.3 Results and Discussion.....	15
2.3.1 The Topology Determination of H3N4.....	18
2.3.2 The Structural Linkage Characterization of H4N2.....	25
2.3.3 Structure Confirmation using Library Matching: H5N3.....	31
2.3.4 H3N3 (m/z 1432) Linkage Characterization by Using Li <sup>+</sup> Adduction Ions .....	36
2.4 Conclusion.....	39
3. GLYCAN BIOMARKER DISCOVERY OF ADULT STEM CELLS USING SEQUENTIAL MASS SPECTROMETRY.....	40

3.1 Introduction.....	40
3.2 Experimental Methods.....	45
3.3 Results and Discussion.....	47
3.3.1 N-glycan MS profile comparison.....	47
3.3.2 The discovery of the sample-specific glycan isomers.....	52
3.3.3 The exploration of N-glycan branching patterns.....	57
3.4 Conclusion.....	64
APPENDIX Isomeric Structures Determination by MS <sup>n</sup> .....	66
REFERENCES.....	72

## LIST OF TABLES

Table 2.1 The overall comparison of the N-glycans isomers identified from chicken ovalbumin using different analytical techniques .....	17
Table 2.2 H3N4 Isomers determined by MS <sup>n</sup> .....	19
Table 2.3 Structural linkages of H4N2 and pathways determined by MS <sup>n</sup> .....	31
Table 2.4 H5N3 isomers determined by MS <sup>n</sup> .....	34
Table 3.1 The distribution of major N-glycans detected MS profiles.....	49
Table 3.2 Side by side comparison of MS <sup>n</sup> spectra obtained from three cell lines within the ion m/z 1810.6+ (H4N3F1) .....	55
Table 3.3: The distribution of the isomers identified within the ion m/z 1810.6+ (H4N3F1) .....	56
Table 3.4 A comparison of structural isomers distribution (T2:T1) .....	60
Table 3.5 Comparison of structural isomers (T-3 vs. T-1+2) distribution.....	62

## LIST OF FIGURES

Figure 1.1 Common Types of N-glycans .....	2
Figure 2.1 The process flow to determine the isomeric structures.....	15
Figure 2.2 Ovalbumin N-glycan ESI MS profile.....	16
Figure 2.3 Nomenclature for carbohydrate fragments.....	18
Figure 2.4 The pathways to determine the topology structures of H3N4 (m/z 1677) .....	21
Figure 2.5 Topology A of H3N.....	21
Figure 2.6 A spectrum with the presence of the conflicted ions.....	23
Figure 2.7 Topology-B of H3N4.....	23
Figure 2.8 Spectrum with ion indicates the conflicted structure.....	24
Figure 2.9 Topology-C of H3N4.....	24
Figure 2.10 Topology and pathways of H4N2 with ambiguous linkages.....	25
Figure 2.11 H4N2 structural linkages characterization (1-3 arm).....	28
Figure 2.12 Isomeric structures linkages characterization for 1-6 arm branched H4N2.....	30
Figure 2.13 The substructure determination of one of the H5N3 isomers using FragLib. ....	32
Figure 2.14 Full linkage structure of H5N3 topology-1.....	32
Figure 2.15 The MS <sup>5</sup> spectrum of H5N3.....	35
Figure 2.16 H3N3 structures linkages characterization by Li <sup>+</sup> adduction.....	38
Figure 3.1 Asymmetric ASCs self-renewal.....	41



Figure 3.2 Three engineered cell lines schemes.....	43
Figure 3.3: Comparative MS profiles analysis of N-glycans from SYM, ASYM and p53SYM .....	48
Figure 3.4: The group abundance distribution of N-glycans across the three cell lines: SYM, ASYM, and p53SYM.....	50
Figure 3.5: N-glycan major biosynthetic steps that localized in ER and Golgi apparatus of mammalian cells.....	51
Figure 3.6 Structural templates.....	58
Figure 3.7 Comparison of relative abundance of structures T2: T1.....	61
Figure 3.8 Comparison of relative abundance of structures T3: T(1+2) .....	63
Figure 3.9 GlcNAc transferases in N-glycan branching .....	64

## ABSTRACT

### N-LINKED GLYCANS CHARACTERIZATION AND BIOMARKER DISCOVERY BY SEQUENTIAL MASS SPECTROMETRY

by

Hongqin (Jenny) Jiao

University of New Hampshire, September 2009

N-linked glycans from a well-studied glycoprotein, (hen albumin) were characterized using ion trap MS ( $MS^n$ ) without chromatographic separation. From the spectral profile, ten major ion compositions were selected for study, which resulted in the characterization of 37 isomeric structures, 16 of which were previously unreported. These findings were compared with previously reported studies using different analytical strategies: IMMS, MS/MS.  $MS^n$  not only identifies but also structurally characterizes components in the absence of any adjunct instrumentation. Selected examples have been detailed.

A major challenge for adult stem cell (ASC) research is the lack of specific biomarkers for their identification and isolation. In this study, carbohydrates were evaluated as candidate biomarkers that mimic the defining property of ASCs, asymmetric self-renewal. Composition differences, ASC-specific structures and the distribution of structural branching patterns were discovered as ASC-specific biomarkers using ion trap  $MS^n$ .



# CHAPTER I

## INTRODUCTION

### 1.1 Glycans and Glycoproteins

Three major classes of biomolecules are required for a living cell: proteins, nucleic acids and carbohydrates<sup>1</sup>. From DNA to RNA to protein, the gene-template-based synthesis and relatively simple linear connection make protein study fairly straightforward and successful. However, proteins alone can not account for much the biological complexity and it is the post-translated modifications (PTM) that provide diversity of organisms. A more complete account of biological function must consider the major roles played by glycoconjugates, e.g., glycoprotein and glycolipids.

Glycoproteins are glycoconjugates in which oligosaccharide chains (called glycans) are covalently attached to a protein. The glycans are frequently present on the outer surface of the cell and can be easily recognized by other cells or organisms. The attached glycans influence biosynthesis, stability, function and turnover of the conjugate protein. Also the glycans play significant roles themselves by using the protein as a platform for positioning<sup>2, 3</sup>.

In general, proteins exhibit two categories of glycans: N-glycans and O-glycans. An N-glycan is a saccharide chain covalently linked by an N-glycosidic bond to asparagine. The most common N-glycosidic bond is acetylglucosamine

(GlcNAc) connected to an asparagine amide residue of a protein with a consensus peptide sequence: Asn-X-Ser/Thr (X can be any amino acid except proline). All N-glycans share a common pentasaccharide core:  $\text{Man}\alpha 1-6(\text{Man}\alpha 1-3)\text{Man}\beta 1-4\text{GlcNAc}\beta 1-4\text{GlcNAc}\beta 1-\text{Asn-X-Ser/Thr}$ . An O-linked glycan is covalently bounded by an O-glycosidic bond from N-acetylgalactosamine (GalNAc) to the hydroxyl group of serine or threonine. Instead of one common core for N-glycans, there are a series of core structural classes for O-glycans<sup>2</sup>.

## 1.2 N-glycans

It is estimated that at least 50% of human proteins are N-glycosylated<sup>4</sup>. N-glycans can be classified into three groups: high-mannose, hybrid and complex, which are summarized in Figure 1.1.

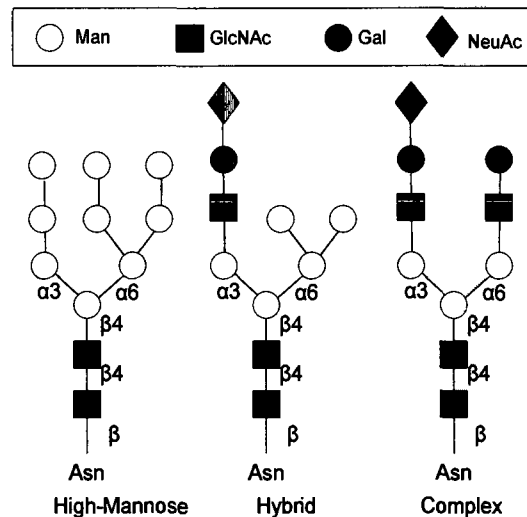


Figure 1.1 Common Types of N-glycans

### N-glycans Biosynthesis

Biosynthesis of eukaryotic N-glycan starts from the endoplasmic reticulum (ER) and finishes in the Golgi apparatus. Compositional and functional diversity

of glycan residues is a distinguishing characteristic of protein glycosylation, which is a consequence of glycosidases and glycosyltransferases<sup>5</sup>. Assembly of N-glycans occurs mainly in four phases: (1) formation of Dolichol-P-P-Glycan precursor -- lipid-linked oligosaccharide (LLO); (2) transfer oligosaccharide to proteins; (3) sugar residues early trimming process; (4) maturation of the glycans. The first two phases have been conserved in all eukaryotic cells. However, the extent of trimming and maturation shows loose genetic relationship<sup>6</sup>.

Glycans are synthesized by a series of glycosyltransferases and degraded by glycosidases. The glycosidases and glycosyltransferase are regulated by gene expression and define a sequence quantity probability but not the exacting sequence. The product of one enzyme reaction can be a preferred substrate for another enzyme. This sequential mechanism can also be substituted by branching. The availability of sugars, the acceptor substrates, the competition among enzymes, the pH levels, the micro environment and so on, all participate in the regulation of a glycan's assembly<sup>2, 7</sup>.

### **Heterogeneity of N-glycoproteins and N-glycans**

Glycoproteins can be highly heterogeneous in several aspects: First, the protein sequence and conformation can cause N-glycoprotein diversity by affecting substrate enzyme affinity. Secondly, glycan-specific sites can be another reason of heterogeneity. If there are more than one sequon of Asn-X-

Ser/Thr on the protein, each sequon could produce a subset of glycans (called glycoforms), and hence glycoproteins. A sequon is only a potential glycosylation site of N-glycans, glycosylation depends on many other factors. Thirdly, glycoforms on a specific site dramatically enlarge the glycoprotein. The glycans are not directly encoded by the genome and their variable structures are highly diversified. Therefore, one protein when glycosylated may exhibit a large heterogeneous pool.

The diversity of N-glycans can be initiated at the monomer level. At least ten monomers are recognized, and the number is growing. D-xylose (Xyl), D-glucose (Glc), D-galactose (Gal), and D-mannose (Man), *N*-acetyl-D-glucosamine (GlcNAc) and *N*-acetyl-D-galactosamine (GalNAc), L-fucose (Fuc), D-glucuronic acid (GlcA) and L-iduronic acid (IdoA), *N*-acetylneuraminic acid *N*-glycolylneuraminic acid are found in the cells of “higher” animals. Also, each monomer may exhibit modifications, such as esterification, acetylation and methylation on different positions, making the glycan pool even larger. The linkage between monosaccharides may be stereocally different referred to as anomeric types ( $\alpha$  or  $\beta$ ). In addition, each monomer can form glycosidic bonds with other monomers, referred to as branching, a common feature of glycosylation.

### **Importance of N-linked Glycosylation**

A single protein can give rise to numerous glycoproteins, which may function in variable ways in tissues. In vivo studies have indicated that changes

of sialylation and branching can affect erythropoietin (EPO)'s stability, half-life, and activity. N-linked glycosylation functions can be probed by using inhibitors of glycosylation, or gene mutations<sup>8,9</sup>. During disease states glycan profile patterns are known to change. This new pattern may affect the biological activity with potential adverse consequences. Therefore, these studies not only provide an understanding of glycan structure, but also reveal that certain glycosyltransferases may be targets for effecting diseases therapeutics<sup>10-12</sup>.

N-glycosylation controls of glycoprotein folding. The consequences of altered glycosylation can be devastating<sup>13</sup>. Many N-glycan structural differences have been reported to be related to the process of preparing cell specific structures<sup>12, 13</sup>. To identify the biological consequences of glycosylation, a complete glycan structural characterization is required. Because of the complexity, this has been a challenging endeavor.

### **1.3 Analytical tools for glycan structures**

A variety of analytical technologies have been developed for characterizing carbohydrate structures. NMR has been considered as the most comprehensive tool for *de novo* glycan characterization. However, the instrument sensitivity, molecular weight limitations, the requirement of relatively pure sample amounts, and spectra interpretation complexity make the practical applications of NMR quite limited<sup>14</sup>. Various chromatographic and electrophoretic techniques



have also been utilized for the separation and identification of glycans<sup>15-18</sup>. Because of their high sensitivity, simplicity and specificity, mass spectrometry-based methods have gained popularity in characterizing glycan structures.

### **Tandem Mass Spectrometry (MS/MS)**

Due to its great success in the field of proteomics, MS/MS based approaches have attracted much attention in the analytical glycomics field. Researchers have applied the protocols directly from proteomics. By using software tools to correlate glycan MS/MS fragment spectra to a database, one can identify a plausible structure. Nevertheless, the MS/MS spectra are inadequate to define the exacting details of a carbohydrate structure. The MS/MS based methods have multiple problems differentiating glycan structures. But clearly, an MS/MS spectrum can not contain adequate information to assign a glycan structure unambiguously<sup>19</sup>.

### **Ion Trap Sequential Mass Spectrometry (MS<sup>n</sup>)**

In an attempt to overcome the limitations outlined above, an ion trap MS<sup>n</sup> approach has been developed in the Reinhold lab at University of New Hampshire for the past several years and more recently, integrated with software tools to handle the plethora of data<sup>20-23</sup>. With an ion trap instrument, MS<sup>n</sup> allows molecules to be disassembled sequentially: selected precursor ions are isolated in the gas phase and further fragmented. The multiple step precursor-product

relationships ( $MS^n$  pathways) define components of structure. By re-tracing the pathway, the topology can be defined by using simple logical reasoning.

A common work flow is outlined in Figure 2.1 in chapter II. Two distinct stages are involved during the process: (1) Determination of isomeric topology; and (2) Identification of inter-residue linkages. By selecting signature ions at each step, the product spectra provide information for major glycan topology determination. Abundant ions that fail to account for the perceived structures indicate the presence of additional isomers. These aberrant ions are then selected for further study<sup>23</sup>. Once the major topologies are determined, the linkage information can be characterized by looking for cross-ring cleavage fragments, and spectral matching in the fragment library<sup>21</sup>.

### **Importance of Sample Derivatization**

Derivatization has been the mainstay of oligosaccharide structural investigation. Methylation<sup>24</sup> is an effective way of blocking the hydroxyl groups to position monomer fragments in a glycan array and additionally eases the cleanup process by lipophilic extraction. Methylation also enhances ion signals. In summary, permethylation is an important component for carbohydrate structure determination.<sup>20</sup>

### **Brief Overview of the Thesis Topics**

In this thesis, two biological application examples are described in the chapter II and III to exemplify our overall strategy for glycan structural characterization

using the MS<sup>n</sup> technology. The two examples address the typical challenges facing glycoprotein analysts, including how to resolve glycan isomeric structure within one real biological sample (Chapter II: Isomer Characterization of N-linked Glycans from Chicken Ovalbumin by Ion Trap MS<sup>n</sup>), and how to perform the comprehensive multiple sample comparison analysis in order to identify potential glycan biomarkers (Chapter III: Glycan Biomarker Discovery of Adult Stem Cells Using Sequential Mass Spectrometry).

## CHAPTER II

# ISOMER CHARACTERIZATION OF N-LINKED GLYCANS FROM CHICKEN OVALBUMIN BY ION TRAP MS

### 2.1 Introduction

The diversity of glycan biological roles has been studied for decades. Variations have been found with development, differentiation, malignancy, and phylogeny of animals. However, glycan structural characterization has been a formidable analytical task hindering biomedical research. An integrated analytical/computational strategy based on ion trap MS<sup>n</sup> technology has been under development in the Reinhold lab in the past several years to attack this challenging problem. This chapter contrasts the results from this method with the reported findings of other laboratories using a widely accepted biological reference standard, the N-glycans from chicken ovalbumin.

Chicken ovalbumin comprises 60 - 65% of the total protein in egg white and is one of the first proteins isolated in a pure form<sup>25</sup>. It was defined as a glycoprotein more than seventy years ago<sup>26</sup>. At least thirty glycan compositions have been detected from chicken ovalbumin on its single glycosylation site. The different glycoforms share a common core structure: Man $\alpha$ 1-6(Man $\alpha$ 1-3)Man $\beta$ -4GlcNAc $\beta$ 1-4GlcNAc -Asn293<sup>27-30</sup>. Chicken ovalbumin has been used as a

standard glycoprotein for the purpose of new analytical methods development and validation for decades.

## **2.2 Experimental Methods**

### **N-linked glycans release**

Chicken egg ovalbumin (grade V) was obtained from Sigma-Aldrich (St. Louis, MO). Peptide-N-glycosidase F (PNGase F) deglycosylation kits were purchased from New England Biolabs Inc. (Ipswich, MA). 1 mg ovalbumin was dissolved in 100  $\mu$ L of 1X denaturing buffer (10 $\mu$ L denaturing buffer 10x concentrated with 90  $\mu$ L HPLC water) and denatured by heating at 98°C for 20 minutes. The sample was cooled to room temperature and to it was added 10 $\mu$ L 10X G7 buffer (0.5 M sodium phosphate, pH=7.5), 10  $\mu$ L NP-40, and 3 $\mu$ L PNGase F enzyme. Sample was vortexed (~10 seconds), centrifuged for 1 minute and then placed in an oven (37°C) for overnight incubation.

### **N-linked glycan separation**

The enzymatically released N-linked glycan solution was dissolved in 5% (v/v) HPLC grade Acetonitrile (ACN), 95% (v/v) HPLC grade water, and 0.1% trifluoroacetic acid (TFA). Proteins were removed by solid phase extraction with a C18 column (SPE) (Alltech Associates, Deerfield, IL). The column was pre-wetted with 5 mL of the following solutions in the sequence: (1) 100% HPLC

grade methanol; and (2) A solution of 5% (v/v) HPLC grade ACN, 95% (v/v) HPLC grade water and 0.1% TFA. Then, the sample was loaded onto the column and washed with 5 mL of 5% (v/v) HPLC grade ACN, 95% (v/v) HPLC grade water and 0.1% TFA. The eluent was immediately collected and dried in a SpeedVac concentrator (Savant Instruments, Holbrook, NY).

### **N-linked glycan purification**

A porous graphitized carbon (PGC) retentive solid phase extraction column was used for additional purification and desalting of the N-linked glycan sample. The column was prepared with the procedure as described<sup>32</sup> with slight modifications. Dry PGC (Carbograph 120/400, Alltech Associates, Deerfield, IL) was washed three times with 4 mL of the following solutions: (1) 1M sodium hydroxide; (2) HPLC grade water; (3) 80% aqueous acetonitrile (ACN) with 0.1% Trifluoroacetic acid (TFA); (4) 25% aqueous ACN with 0.05% TFA; (5) 25% aqueous ACN; (6) HPLC grade water. The tube was vortexed, centrifuged and the supernatant was removed from each washing. The prepared PGC was loaded to a 1.5 mL reservoir column and conditioned with 3 mL HPLC grade water. Dried glycan sample was dissolved in 1 mL HPLC water and loaded onto the column. Salts were removed by 5 mL water wash and neutral glycans were collected by elution with 1 mL 25% (v/v) acetonitrile aqueous solution 5 times. The 5 mL eluent was dried in a SpeedVac concentrator.

### **N-linked glycan reduction**

N-glycans were reduced to alditols with a 500  $\mu$ L solution of sodium borohydride ( $\text{NaBH}_4$ ) (10mg/mL in 0.01 M NaOH) and sitting for 1 hour in ice followed by leaving overnight at room temperature. The reaction was terminated by drop-wise addition of glacial acetic acid. Sample was diluted with 3mL ethanol and dried with a stream of nitrogen. Borates were removed by esterification with methanol sequentially adding and drying 3mL of 1% acetic acid methanol solution 3 times and then by 3mL of toluene for 3 times. Dried samples were dissolved in water and desalted on a DOWEX AG 50 W X8-400 cation-exchange resin (Sigma-Aldrich, St. Louis, MO). The resin was washed by 5 mL HPLC grade water 10 times and loaded into a 1.5 mL reservoir. Glycans were eluted by 5mL water and dried with SpeedVac.

### **N-linked glycan permethylation**

Permethylation<sup>24</sup> was carried out by dissolving the dried sample in 500  $\mu$ L dimethyl sulfoxide (DMSO) (Sigma-Aldrich, St. Louis, MO). Sodium hydroxide powder (99.99%) was added to the glycan-DMSO solution. After 30 minutes of vortexing, to produce a suspension, 100  $\mu$ L of methyl iodide was added to perform the permethylation. The reaction tube was wrapped with parafilm and aluminum foil, and then vortexed for 1 hour. Two milliliters of water was added to the reaction tube in ice bath to terminate the reaction. Permethyated glycans were extracted three times with 1 mL of dichloromethane (HPLC grade, Fisher Scientific, Fair Lawn, NJ). The extract was washed 5 times with 2 mL of HPLC

water. The methylated glycans were dried under nitrogen stream, and analyzed by MS.

### **MALDI-TOF Mass Spectrometry**

Released, reduced and permethylated glycans were analyzed with a MALDI-CFR mass spectrometer (Krotos-Shimadzu Analytical, Manchester, UK) equipped with an ultraviolet 337 nm wavelength nitrogen laser. For native and reduced glycans, the sample was dissolved in 100  $\mu$ L of HPLC grade water, and for permethylated samples, 75% methanol aqueous solution was used. One  $\mu$ L of sample was spotted on the MALDI target plate with 1  $\mu$ L of 2, 5-dihydroxy benzoic acid matrix (DHB) (10 mg/mL in 50% (v/v) acetonitrile aqueous solution). After drying, 0.4  $\mu$ L ethanol (HPLC grade) was applied to the spot for constant signal. Parameters (such as: power, shots accumulated per profile, optimized mass number) were set according to set protocols. Spectra were processed by Kratos Lauchpad software.

### **ESI Mass Spectrometry**

$MS^n$  mass spectra were obtained with a LTQ mass spectrometer (Thermo Fisher Scientific, Waltham, MA) equipped with a TriVersa Nanomate (Advion) automated ion source. Samples were infused at flow rates ranging from 0.30 to 0.60  $\mu$ L/min, and spectra were collected by using Xcalibur 1.4 and 2.0# software (Thermo Fisher Scientific). Signal averaging was accomplished by 5 micro-scans within each scan and adjusted 50 to 300 scans in each spectrum according to



the level of signal to noise. This improved the spectral quality and was necessary at the higher orders of  $MS^n$  due to diminishing signal, particularly beyond  $MS^5$  or  $MS^6$ . Collision parameters were left at default values with normalized collision energy set to 35% or to a value leaving a minimal precursor ion peak. Activation Q was set at 0.25 and activation time for 30 ms. Changing collision energy did not affect the relative abundance of product ions, but change the precursor ion abundance. The Activation Q and activation time were kept constant since they can affect the product ion abundances<sup>20</sup>.

### **Data Analysis**

In-house bioinformatics tools (Glycoscreen, CompositionFinder et al) developed by Hailong Zhang were used to assist data analysis. A schematic flow process used to determine structures of N-glycans has been outlined, (Fig. 2.1). Initially, the ion compositions of N-glycans were assigned using Glycoscreen. Then, selected ions from each  $MS^n$  spectrum were considered to determine topologies. During this process, each spectrum was carefully examined to ensure that there no conflicting ions. If there were, then these aberrant ions were selected resolve the different topologies. Once the topologies were resolved, the characteristic fragments resulting from cross-ring cleavages within small oligomers were compared with fragment library standards to help define linkages.

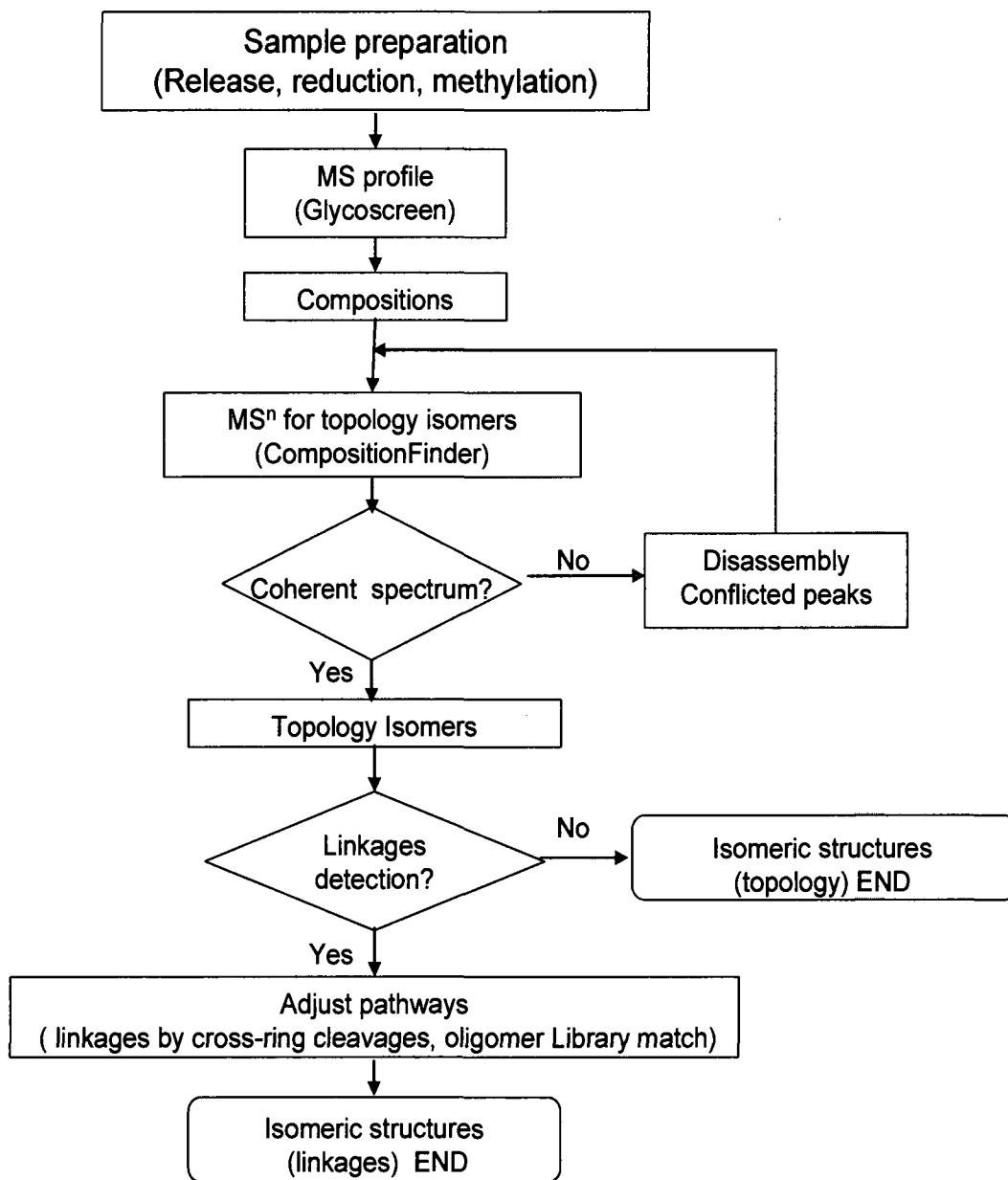


Figure 2.1 The flow process to determine structures.

## 2.3 Results and Discussion

### MS profile and N-glycan compositions

The ESI MS profile (Fig. 2.2) is a spectrum of the reduced, methylated N-glycans from chicken ovalbumin with the 10 most abundant ions labeled with  $m/z$  values and their corresponding compositions. The N-glycans share a common core of  $\text{Man}\alpha 1\text{-6}(\text{Man}\alpha 1\text{-3})\text{Man}\beta\text{-4GlcNAc}\beta\text{1-4GlcNAc}$ . The compositions of the 10 most abundant ions in MS profile are: H3N2, H4N2, H3N3, H5N2, H4N3, H3N4, H6N2, H5N3, H4N4 and H3N5. H indicates hexose and N indicates HexNAc. These compositions were selected for  $\text{MS}^n$  analysis.

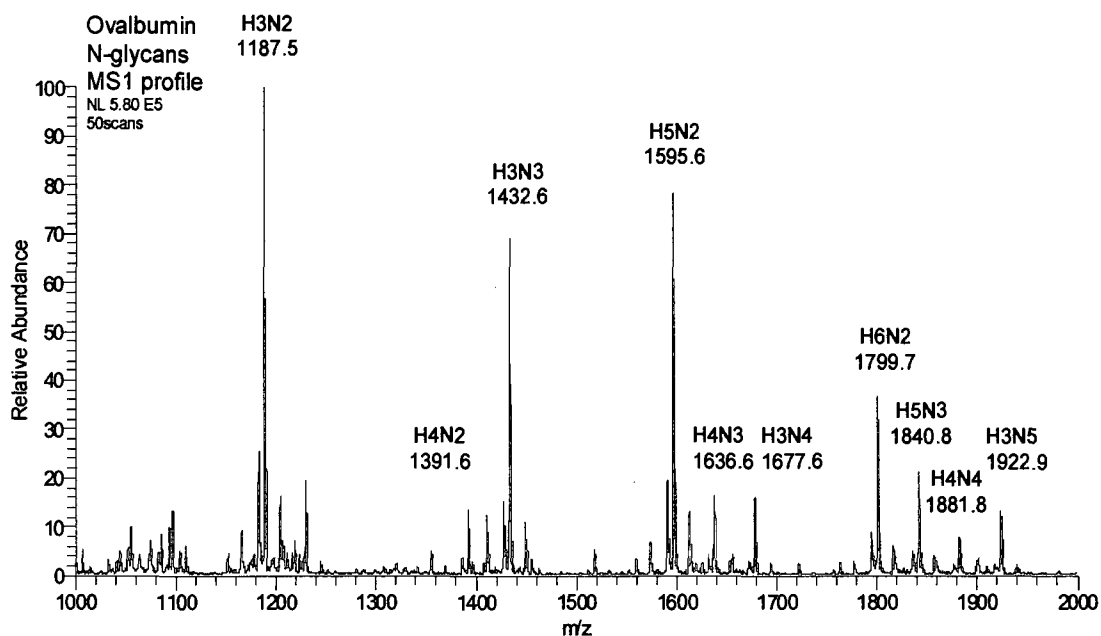


Figure 2.2 The ESI MS profile of N-linked glycans from chicken ovalbumin with the 10 most abundant ions labeled with  $m/z$  values and their corresponding compositions (H: Hexose; N: HexNAc).

## **Isomer characterization using sequential mass spectrometry**

In total, 36 N-glycan isomers have been identified within the 10 selected compositions. 15 of which were structures previously unreported. The complete list of isomers and supporting evidence has been compiled in Appendix. Due to the large volume of information and the thesis page limit, only a few selected examples are presented in the following sections to exemplify the general procedures and techniques which have been employed during this study.

Glycan Composition	m/z (reduced; permethylated; Na <sup>+</sup> )	# of isomers IMS/MS*	# of isomers MS/MS**	# of isomers MS <sup>n</sup>
H3N2	1187.5	1	1	1
H4N2	1391.6	1	1	3
H3N3	1432.6	4	3	3
H5N2	1595.6	1	1	3
H4N3	1636.6	2	2	6
H3N4	1677.6	1	2	4
H6N2	1799.7	2	1	4
H5N3	1840.8	3	2	4
H4N4	1881.8	4	3	5
H3N5	1922.9	1	2	3
Total		20	18	36

Table 2.1: The overall comparison of the N-glycans isomers identified from chicken ovalbumin using different analytical techniques. Reference: Clemmer et al. \* (2008); Reference: Lattova and Perreault \*\* (2004); Harvey et al. \*\* (2000).

The overall comparison of the isomers from chicken ovalbumin using different analytical techniques<sup>#19, #31</sup> is shown in Table 2.1. In general, the sequential mass spectrometry-based approach provides a more comprehensive

evaluation of isomers than either IMS/MS (ion mobility mass spectrometry) or MS/MS methods. The nomenclature of carbohydrate fragmentation are shown in Figure 2.3<sup>33</sup>.

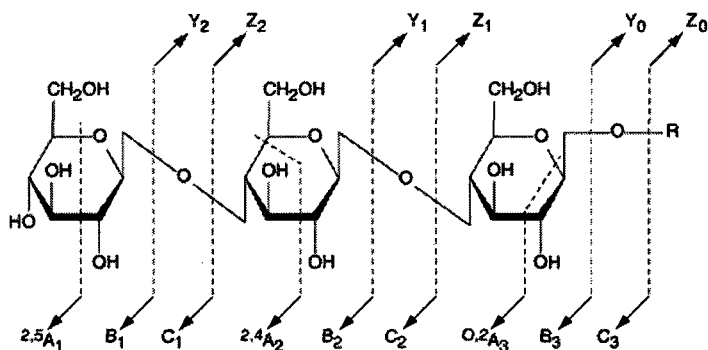


Figure 2.3 Nomenclature for carbohydrate fragments. This figure is modified from Domon and Costello (1988).

### **2.3.1 The Topology Determination of H3N4 (m/z 1677)**

The process of determining topology for the ion composition H3N4 (m/z 1677) is detailed to demonstrate the approach to resolving isomers. The diagnostic MS<sup>n</sup> pathways and the determined isomer structures are summarized (Table 2.2), along with the corresponding glycan fragments pathways.

Isomeric structures D, E and F were previously reported. Structure D and E are reported by Lattova et al<sup>31</sup> (MS/MS), while structure E and F are reported by Harvey<sup>28</sup> using multiple techniques (mainly NMR). In this thesis study, structure D was found after linkage elucidation of topology A; isomeric structure E and F were detected after linkage characterization of topology B. The fourth isomeric structure G was revealed after the linkage elucidation of topology C.

Unlike structure D, E or F which has one HexNAc attached to the core mannose, this structure has two HexNAcs attached to one of the 3-linked core mannose.

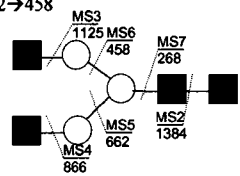
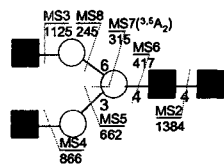
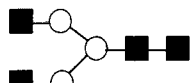
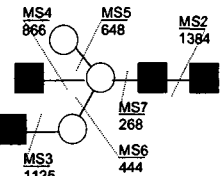
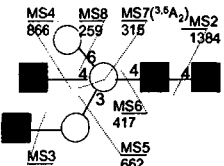
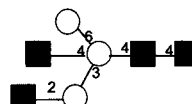
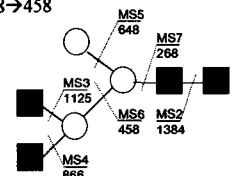
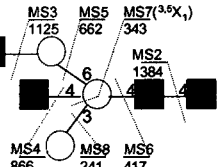
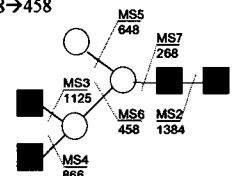
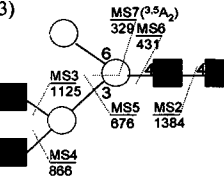
Glycan composition	Isomers by MS <sup>n</sup>		Isomers reported (references)	
	Topology Isomers	Linkage Isomers		
H3N4 m/z 1677.6 (Na <sup>+</sup> )	1677→1384→1125 →866→662→458 →268  <b>A</b>	1677→1384→1125→866→662→417 →315( <sup>3,5</sup> A <sub>2</sub> -ion)→245  <b>D</b>	 <b>D</b> (reference 31)	
	1677→1384→1125 →866→648 →444→268  <b>B</b>	1677→1384→1125→866→662→417 →315( <sup>3,5</sup> A <sub>2</sub> -ion)→259  <b>E</b>		 <b>E</b> (reference 28&31)
	1677→1384→1125 →866→648→458 →268  <b>C</b>	1677→1384→1125→866→662→417 →343( <sup>3,5</sup> X <sub>1</sub> -ion)→241  <b>F</b>		
1677→1384→1125 →866→648→458 →268  <b>C</b>	1677→1384→1125→866→676→431 →329( <sup>3,5</sup> A <sub>2</sub> -ion)(and 213)  <b>G</b>			

Table 2.2 H3N4 isomers determined by MS<sup>n</sup>.

Generally, fragments can be classified into two groups: glycosidic bond cleavages and cross-ring cleavages<sup>34</sup>. Glycosidic cleavage occurs between two monomers, which may yield B-, C-, Y- or Z- type ions. A cross-ring cleavage ruptures two bonds in a monomer, which gives A- and X- type ions. Glycosidic

bond cleavages define a glycan's topology, while cross-ring cleavages are essential in determining of inter-residue linkages. The number of glycosidic bonds in each ion composition provides a clue of other associated residues. For a permethylated sample, glycosidic bond cleavages provide specific pyranosyl-1-ene (B- and Z-type ions) or open hydroxyl (Y- and C-type ions) that would be products of the ruptured bond. These cleavage identifiers are defined as "scars". By selecting compositions rich in such information, monomers can be positioned in a topology array.

As an example, to elucidate the topology of H3N4, the ion with  $m/z$  1677.6 was selected for MS<sup>2</sup> analysis (Fig. 2.2). A product ion ( $m/z$  1384.5) was the fragment residue H3N3 with one scar. This can be explained by a neutral loss of reduced terminal HexNAc (293u) leaving one scar on the product ion. By selecting that fragment ( $m/z$  1384.5) for MS<sup>3</sup>, a dominant ion ( $m/z$  1125.4) was observed fitting the composition H3N2 with two scars. Similarly, an informative ion from each spectrum can be selected; (pathway 1:  $m/z$  1677  $\rightarrow$  1384  $\rightarrow$  1125  $\rightarrow$  866  $\rightarrow$  662  $\rightarrow$  458  $\rightarrow$  268). Considering the known N-glycan core structure, the disassembled products define the H3N4 topology without ambiguity, (Figure 2.5).

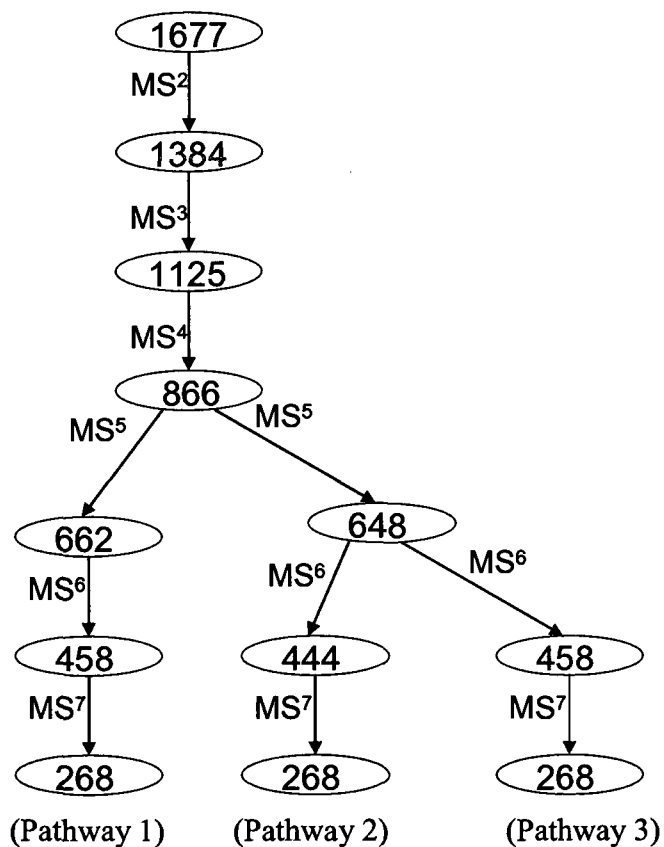


Figure 2.4: The pathways to determine the topology structures of H3N4 (m/z 1677).

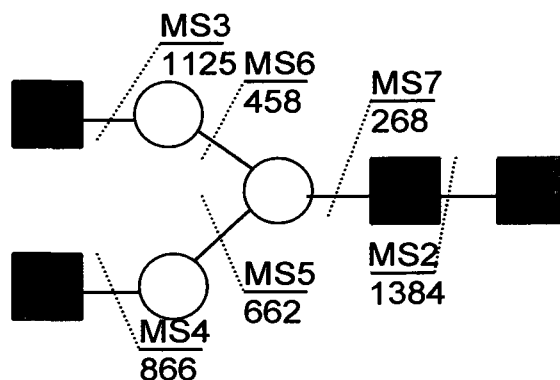


Figure 2.5: Topology A of H3N4: 1677 → 1384 → 1125 → 866 → 662 → 458 → 268.



The ion abundance is not the only factor considered in parent ion selection, composition that includes monomers and scars is another. e.c. structural information is another. Ion fragments inconsistent with a single structure are indicators of isomers. Such ions in the ion trap can be further isolated, activated (CID) and disassembled, (Fig. 2.6). The ion with m/z of 662.3 (Fig. 2.5) represents a major structure. The ion m/z 648.3 in the same spectrum represents the neutral loss of a fully methylated hexose inconsistent with that proposed structure. To examine this ion, pathway 2 was chosen m/z 866 → 648 → 444 → 268. By reevaluating the pathway, it was clear another topology was determined, (Fig. 2.7).

Figure 2.8 shows another spectrum with the presence of isomeric ions. The ion of m/z 444 in MS<sup>6</sup> suggests Topology B (Table 2.2) of H3N4 in MS<sup>5</sup>, but the ion with m/z of 458 in MS<sup>6</sup> indicates an additional isomer. To examine this ion, pathway 3 was selected, m/z 648 (neutral loss: one-scar H; 218u) → 458 (neutral loss: three-scar H; 190u) → 268 (a two-scar N charged with Na<sup>+</sup>) (neutral loss: three-scar H; 190u) was executed. Figure 2.9 shows the topology of H3N4 (Topology C) determined from this pathway.

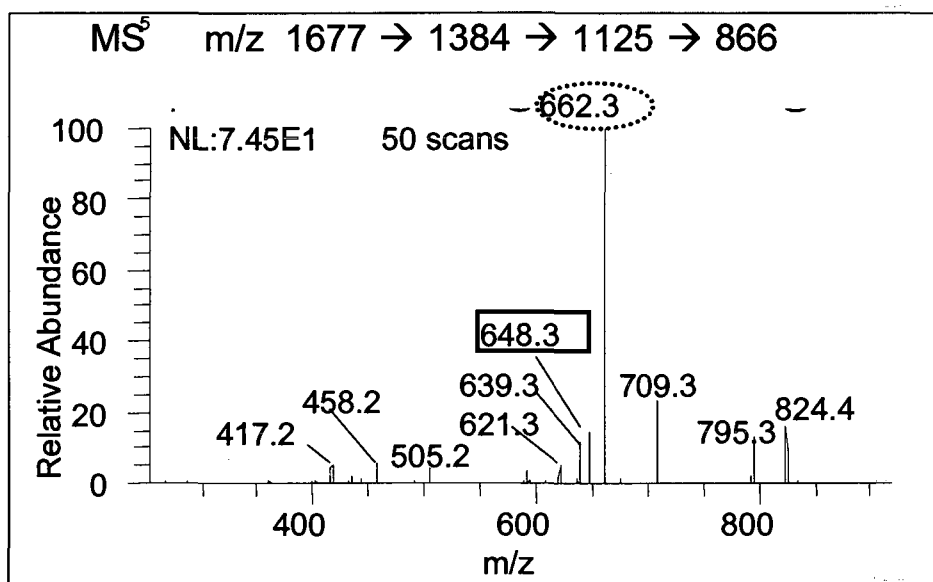


Figure 2.6: A spectrum with the presence of conflicted ions. The ion with m/z of 662.3 represents a major characteristic structure. The ion with m/z of 648.3 conflicts to that major structure.

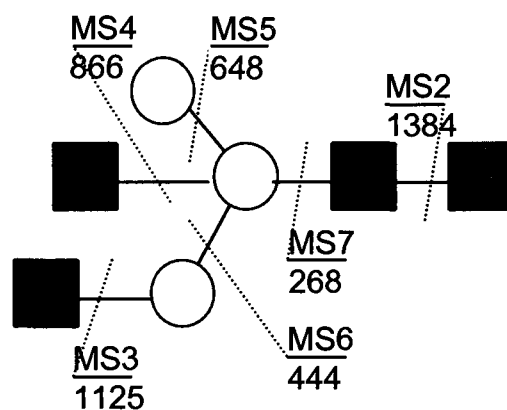


Figure 2.7: Topology B of H3N4: 1677 → 1384 → 1125 → 866 → 648 → 444 → 268.

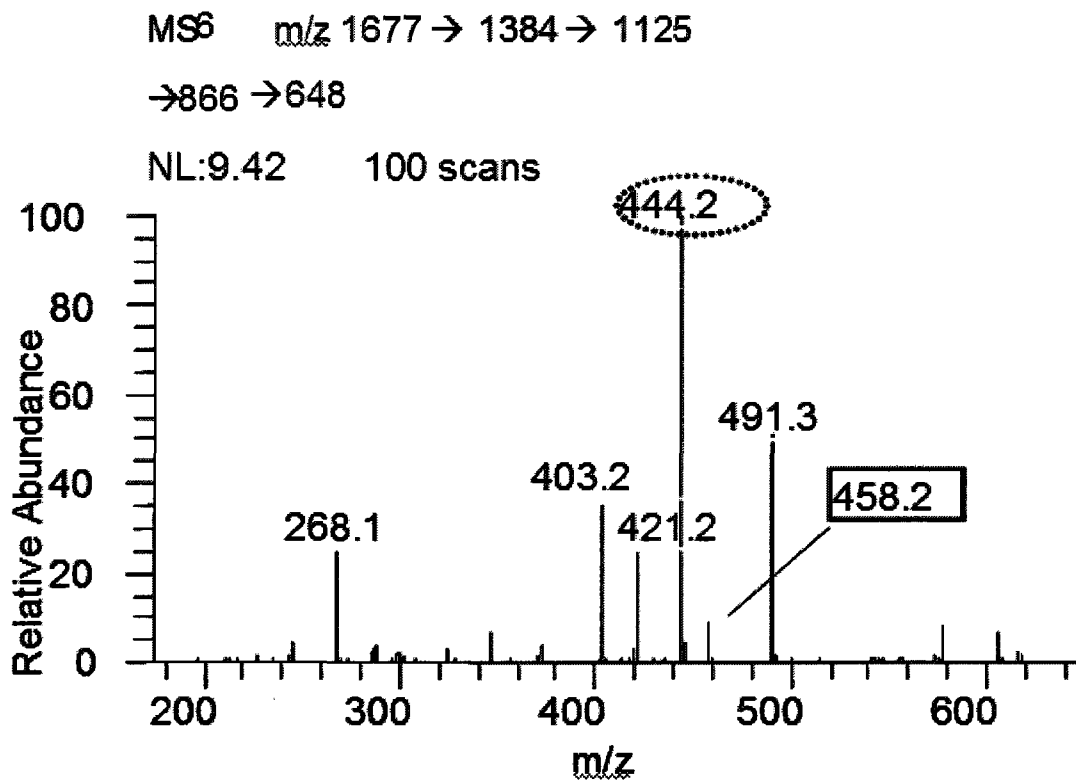


Figure 2.8: A spectrum with ion indicates the existence of the conflicted structure. The ion with  $m/z$  of 444.2 represents Topology B of H3N4. The ion with  $m/z$  of 458.2 represents a new structure.

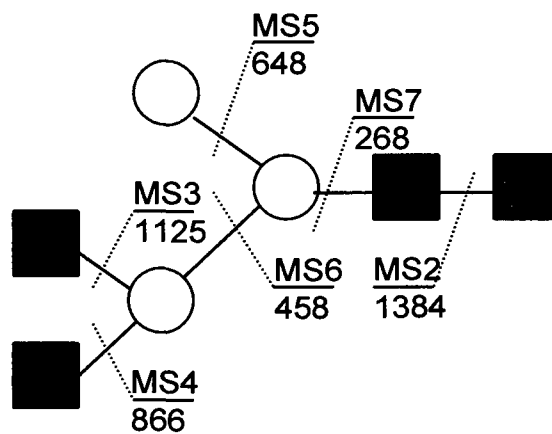


Figure 2.9: Topology C of H3N4: 1677  $\rightarrow$  1384  $\rightarrow$  1125  $\rightarrow$  866  $\rightarrow$  648  $\rightarrow$  458  $\rightarrow$  268.

From MS<sup>2</sup> to MS<sup>4</sup> spectra, the topology A of H3N4 (Table 2.2) is suggested by the dominant ion m/z 662 and m/z 458. The topology B can be determined by the relatively less abundant ion m/z 648 and m/z 444. However, topology C was not revealed until MS<sup>6</sup> by the fragmentation of ion m/z 648.

### 2.3.2 The Structural Linkage Characterization of H4N2 (m/z 1391)

The composition H4N2 has been selected to explain aspects of structural linkage elucidation. Three isomeric structures have been revealed in the study of this composition, Table 2.3.

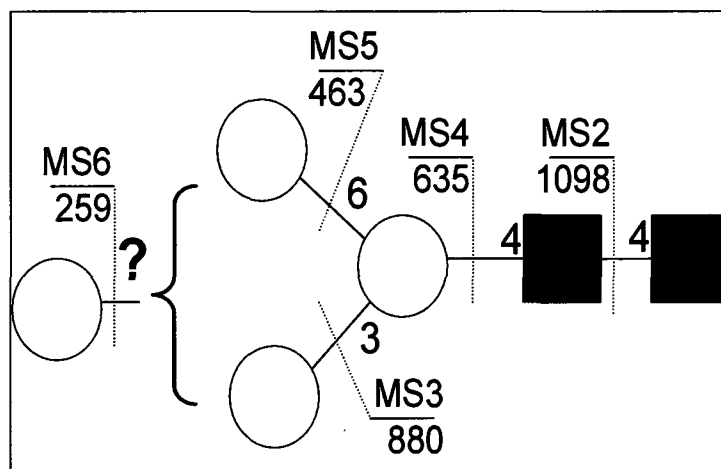


Figure 2.10: Topology and pathways of H4N2 with ambiguous linkages.

The Topology determination pathway is shown in Fig. 2.10. The position and linkage of the terminal hexose (non-core hexose) are unspecified. To determine the topology, the pathway of ion with  $m/z$  1391: (H4N2)  $\rightarrow$  1098 (neutral loss: one-scar R; 293u)  $\rightarrow$  880 (neutral loss: one-scar H; 218u)  $\rightarrow$  635 (neutral loss: two-scar N; 245u)  $\rightarrow$  463 (a one-scar C<sub>2</sub>-ion; charged with Na<sup>+</sup>) (neutral loss: three-scar H; 172u) was followed.

Fragments from glycosidic bond and cross-ring cleavages are both used for the linkages characterization. Figure 2.11 and Figure 2.12 show the procedure of how linkages of H4N2 were characterized. Based on the common core sequence of Man $\alpha$ 1-6(Man $\alpha$ 1-3)Man $\beta$ -4GlcNAc $\beta$ 1-4GlcNAc of N-glycans, it is important to find a characteristic cross-ring cleavage ion in order to determine the position of outer hexose on the core Man $\alpha$ 1-6Man arm or Man $\alpha$ 1-3Man arm. Ions from specific cross-ring cleavages act as an indicator of linkage position<sup>34</sup>. No apparent ion corresponding to a specific linkage was observed in MS<sup>2</sup> or MS<sup>3</sup> spectra (Figure 2.11a and b). However, cross-ring fragments increased upon MS<sup>4</sup> (Fig. 2.11c) providing direct indication of disaccharide linkage. According to previously determined topology, all these cross-ring fragments should come from the inner core Mannose ring rupture linked with a terminal dimer (H2 with 1 scar). In other words, ions with  $m/z$  of 505.3, 533.3, 547.3 and 561.3 are <sup>0,4</sup>A<sub>3</sub>-ion, <sup>3,5</sup>A<sub>3</sub>-ion, <sup>3,5</sup>X<sub>2</sub>-ion and <sup>2,4</sup>A<sub>3</sub>-ion, respectively. These ions were further disassembled and processed for MS<sup>5</sup>.

Dominant ions with  $m/z$  463 in each MS<sup>5</sup> spectrum confirmed the existence of terminal dimer. However, dimers generated by cross-ring cleavages can be

linked to either arm of the core structure, namely,  $\text{Man}\alpha(1-6)\text{Man}$  arm or  $\text{Man}\alpha(1-3)\text{Man}$ . Ions with  $m/z$  547 and  $m/z$  533 were selected as diagnostic ions to further detail linkages.

The ion with  $m/z$  of 547.3 is a  $^{3,5}\text{X}_2$ -ion coupled with a B-type ion, and fragment from the structure specific for defining 1-3 linked antennae. By disassembling this  $\text{X}_2\text{-B}_2$ -ion, the structure of non-core hexose located on the  $\text{Man}\alpha 1-3\text{Man}$  arm is unveiled by  $m/z$  463.2 (one-scar  $\text{C}_2$ -ion) and  $m/z$  445.2 (one-scar  $\text{B}_2$ -ion) in  $\text{MS}^5$  (Fig. 2.11d). Since dominant glycosidic cleavage and weak cross-ring cleavage ions exist in  $\text{MS}^5$  spectrum, the linkage between the outer hexose and  $\text{Man}\alpha 1-3\text{Man}$  core is still not known. The fragmentation  $\text{C}_2$ -ions or  $\text{B}_2$ -ions from this arm is needed.

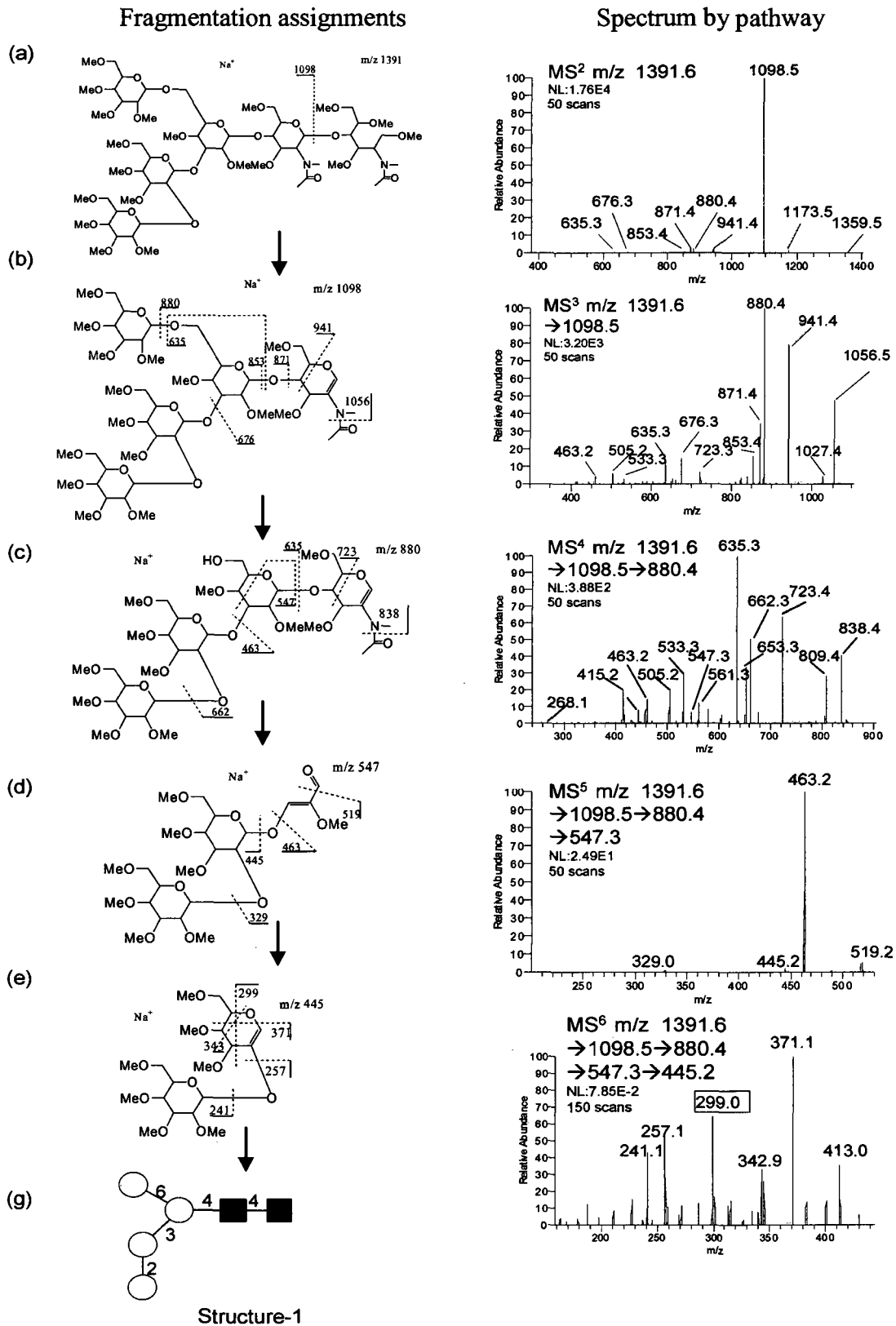


Figure 2.11: H4N2 structural linkages characterization (1-3 arm).

For linkage characterization, usually a B-ion spectrum can provide more information than that of a C-ion. Therefore, instead of dominant C<sub>2</sub>-ion, the B<sub>2</sub>-ion in MS<sup>6</sup> was selected for further disassembly. The dominant peak m/z 371 (Fig. 2.11e) indicates the terminal hexose is not present as the linkage of 1-6. The presence of relatively abundant ion with m/z of 299, which is a major cross-ring cleavage <sup>2,5</sup>X<sub>1</sub>-ion, indicates that the hexose is 1-2 linked to the core Man $\alpha$ (1-3)Man arm. This ion can only be produced by terminal hexose 1-2 linkage (Figure 2.11e).

In MS<sup>4</sup>, the ion with m/z of 533.3 (<sup>3,5</sup>A<sub>3</sub>-ion) is inconsistent with the structure of non-core hexose located on the core Man $\alpha$ (1-3)Man arm (Fig. 2.11 d to f), but fits the structure of Fig. 2.12a. The <sup>3,5</sup>A<sub>3</sub>-ion of m/z 533.3 indicates that the non-core hexose is located on the core Man  $\alpha$ (1-6)Man arm. The <sup>3,5</sup>A<sub>3</sub> ion m/z 533.3 was therefore selected for MS<sup>5</sup>. The C<sub>2</sub>-ion m/z 463.2 and B<sub>2</sub>-ion m/z 445.1 in MS<sup>5</sup> indicate the existence of the outer Hexose linked to the Man  $\alpha$ (1-6)Man arm (Fig. 2.12b). To identify the linkage of the hexose with core Man $\alpha$ (1-6)Man, the B<sub>2</sub>-ion was selected for MS<sup>6</sup>. In the MS<sup>6</sup> spectrum (Fig. 2.12c), the dominant cleavage of m/z 371, a <sup>0,4</sup>X<sub>1</sub>-ion, indicates the absence of any 1-6 linkage between monomers. A major peak of m/z 343, a <sup>3,5</sup>X<sub>1</sub>-ion, could come from the structure with 1-2/3 linkage. By comparing the spectrum in Figure 2.11e, a 1-2 linkage specific ion m/z 299 does not exist in Figure 2.12c. It is believed that a structure with 1-3 linkage should be a major linkage, (Fig. 2.12d). Moreover, the cross-ring cleavage m/z 329 shows the presence of the linkage 1-4/6, which conflicts to the 1-3 linkage structure. Even though at a relatively low



abundance, the peak of  $m/z$  301 still indicates 1-6 linkage structure, and it is believed that Structure 3 exists (Figure 2.12d).

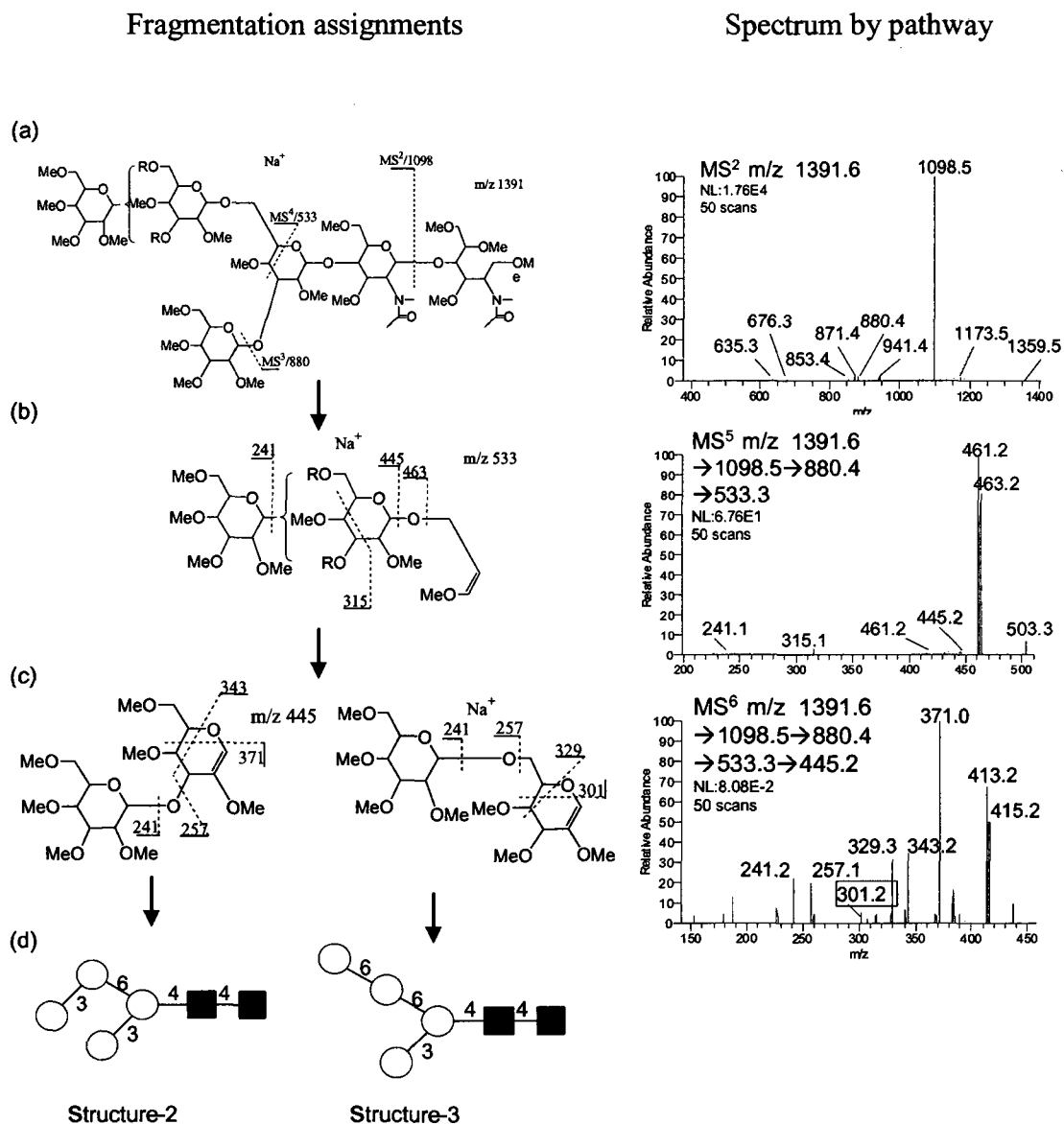


Figure 2.12 Structural linkage characterization for the 1-6 arm branched composition H4N2.

In this way, a total of 3 isomers were identified for H4N2. The isomeric Structure 1 of H4N2 has an outer hexose 1-2 linkage to Man  $\alpha$ (1-3)Man arm of N-glycan core. The isomeric Structure 2 of H4N2 has an outer hexose 1-3 linkage

to Man  $\alpha(1-6)$ Man arm of N-glycan core. The isomeric Structure 3 of H4N2 has an outer hexose 1-6 linkage to Man  $\alpha(1-6)$ Man arm of N-glycan core. Table 2.3 lists the structural linkages and corresponding pathways.

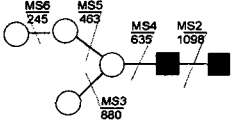
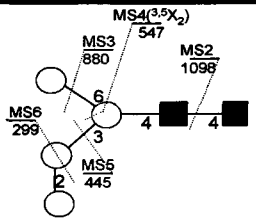
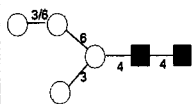
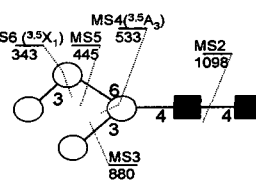
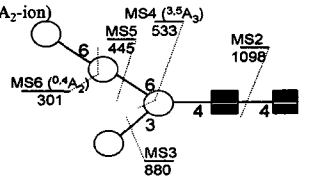
Glycan composition	Isomers by MSn		Isomers reported (references)
	Topology Isomers	Linkage Isomers	
H4N2 m/z 1391.6 (Na <sup>+</sup> )	 <p>1391 → 1098 → 880 → 635 → 463 → 245</p>	<p>1391 → 1098 → 880 → 547 (<sup>3,5</sup>X<sub>2</sub>-ion) → 445 → 299 (<sup>2,5</sup>X<sub>1</sub>-ion)</p> 	 <p>(reference 28)</p>
		<p>1391 → 1098 → 880 → 533 (<sup>3,5</sup>A<sub>3</sub>-ion) → 445 → 343 (<sup>3,5</sup>X<sub>1</sub>-ion)</p> 	
		<p>1391 → 1098 → 880 → 533 (<sup>3,5</sup>A<sub>3</sub>-ion) → 445 → 301 (<sup>0,4</sup>A<sub>2</sub>-ion)</p> 	

Table 2.3 Structural linkages of H4N2 and their pathways determined by MS<sup>n</sup>.

### 2.3.3 Structure Confirmation using Library Matching: H5N3 (m/z 1840. 8)

MSn spectra of carbohydrate samples can be utilized as an exacting structure fingerprint. Structural details, including monomer identification, branching, linkage and stereochemical features can be determined by a combination of glycan disassembly and spectral matching of the products. The library (FragLib) has been generated from commercial samples (small oligomers) and those prepared from well-characterized glycoproteins. In this study, the

spectral library matching method has been used for the determination of carbohydrate substructures.

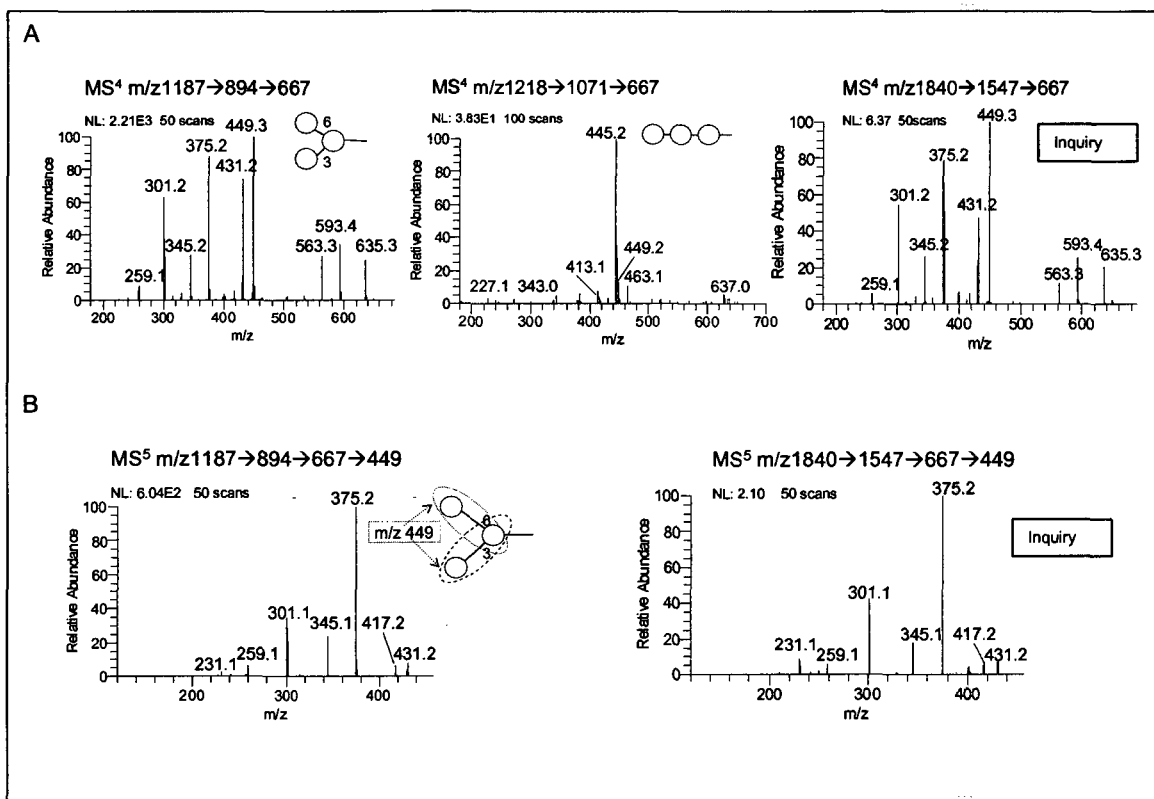


Figure 2.13 The substructure determination of one of the H5N3 isomers using FragLib. A: the comparison of the standard spectra (A-1: branched H3 and A-2: linear) and the MS4 spectrum 1840 → 1547 → 667 of H5N3 (A-3); B: the comparison of the standard spectrum (B-2) and the MS5 spectrum 1840 → 1547 → 667 → 449 of H5N3 (B-2).

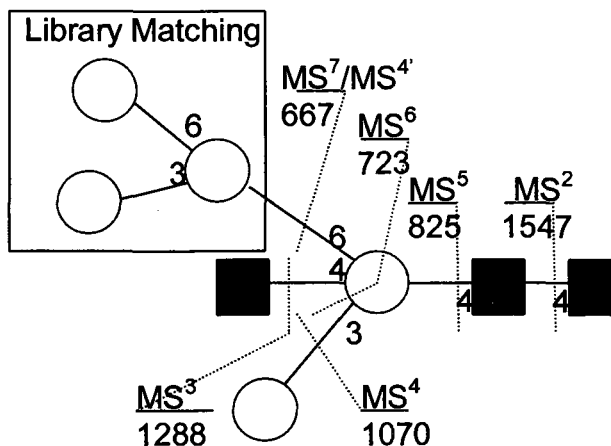


Figure 2.14 A full structural linkage of H5N3 Topology revealed by MS<sup>n</sup> and library matching.

Figure 2.13 illustrates the process of substructure determination of one of the H5N3 isomers using FragLib. The MS4 spectrum (1840 → 1547 → 667) of H5N3 (Fig. 2.16 A-3), which corresponds to a composition as H3 plus one single scar, matches very well with the reference spectrum from the branched standard (Fig. 2.13 A-1), but shows no similarity to the spectrum from the linear standard (Fig. 2.13 A-2). The further comparison of the MS5 spectrum also confirms the results (Fig. 2.13 B).

By using library matching, it is concluded that the fragment of  $m/z$  667.3 from H5N3 has a single structure. This structure is the same as the tri-mannose part,  $\text{Man}\alpha(1-6)\text{Man}\alpha(1-3)\text{Man}$ , of N-glycan core. Fig. 2.14 shows a complete characterized structure after all separated fragments from the parent ion are regrouped.

### **A Comparative Assessment: H5N3 ( $m/z$ 1840.8)**

For the composition H5N3 with  $m/z$  1840.8, a total of 4 topologies were determined by  $\text{MS}^n$  (Table 2.4). Two of them were reported previously (by NMR and MS/MS). They are also included in Table 2.4 for easy comparison. The two isomers defined by the diagnostic pathways of 1840 → 1547 → 1288 → 1070 → 852 → 634 → 444 and 1840 → 1547 → 880 → 635 → 431 are unreported structures.

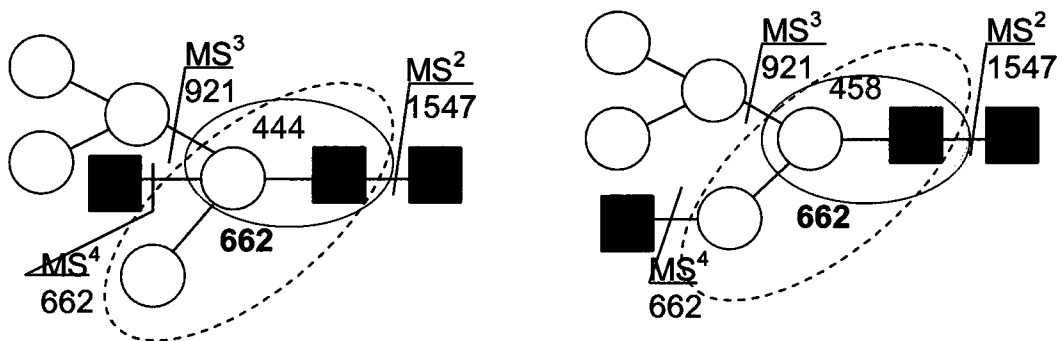
Glycan composition	Isomers by MS <sup>n</sup>		Isomers reported (references)
	Topology Isomers	Linkage Isomers	
H5N3 m/z 1840.8 (Na <sup>+</sup> )	1840→1547→1288 1070→852→634→ 444→268 	1840→1547→ $\left\{ \begin{array}{l} 1288 \rightarrow 1070 \rightarrow 825 \rightarrow 723^{(3,5 A_3)} \rightarrow 667 \\ 667 \rightarrow 449 \text{ (library data matching)} \end{array} \right.$ 	
	1840→1547→ $\left\{ \begin{array}{l} 921 \rightarrow 662 \rightarrow 458 \rightarrow 268 \\ 667 \rightarrow 449 \rightarrow 259 \end{array} \right.$ 	1840→1547→ $\left\{ \begin{array}{l} 1288 \rightarrow 1084 \rightarrow 839 \rightarrow 737^{(3,5 A_3)} \rightarrow 667 \\ 667 \rightarrow 449 \text{ (library data matching)} \end{array} \right.$ 	 (reference 28&31)
	1840→1547→ $\left\{ \begin{array}{l} 1084 \rightarrow 880 \rightarrow 635 \rightarrow 463 \\ 486 \rightarrow 268 \end{array} \right.$ 		 (reference 31)
	1840→1547→ $\left\{ \begin{array}{l} 880 \rightarrow 635 \rightarrow 431 \\ 690 \rightarrow 463 \end{array} \right.$ 		

Table 2.4 H5N3 isomers determined by MS<sup>n</sup>.

(A)

Isomer # 1 of H5N3 (m/z 1810.6)

Isomer # 2 of H5N3 (m/z 1810.6)



(B)

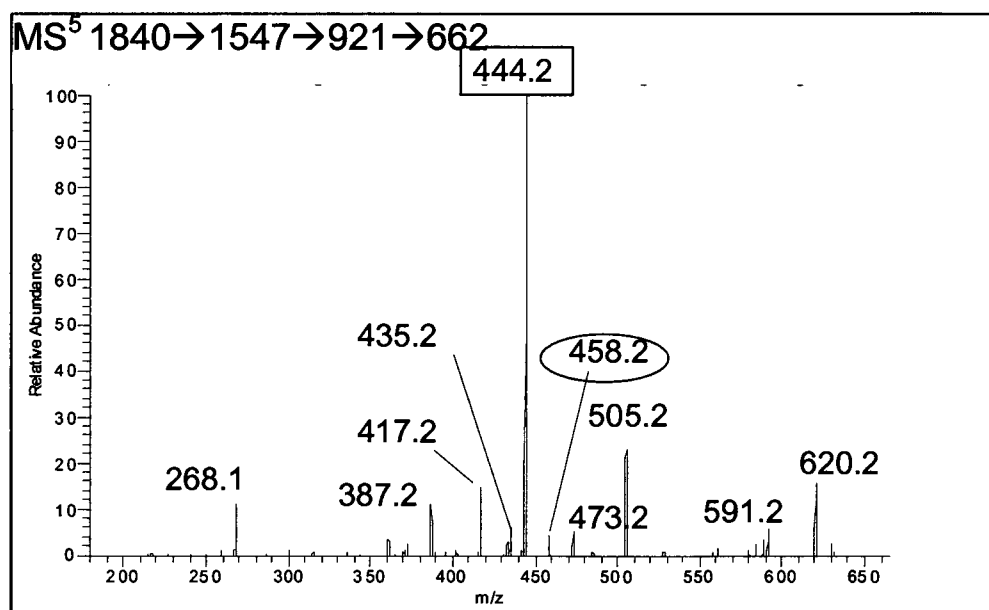


Figure 2.15 The MS<sup>5</sup> spectrum (1840 → 1574 → 921 → 662) of H5N3 provides strong evidence to support the presence of the topology isomer #1, which is missing from the literature. (A) The ions m/z 444 and m/z 458 are the diagnostic ions of H5N3 isomeric structures #1 and #2 respectively. (B) The MS<sup>5</sup> spectrum (1840 → 1574 → 921 → 662) of H5N3

The MS<sup>5</sup> spectrum (1840 → 1574 → 921 → 662) of H5N3 provides strong evidence to support the presence of the topology isomer #1, which is missing from the literature (Figure 2.15). The residue of m/z 662 is H2N1 with 3 scars. The ions with m/z 444 and m/z 458 are the structure-specific ions to the isomer #1 and Isomer #2 respectively. In the MS<sup>5</sup> spectrum (Figure 2.15 B), the ion m/z 444 (H1N1 with 4 scars) is much more abundant than the ion m/z 458 (H1N1 with 3 scars). Therefore, our sequential mass spectrometry-based approach is capable to detect and resolve isomers which might be invisible to the others methods.

#### **2.3.4 H3N3 (m/z 1432) Linkage Characterization by Using Li<sup>+</sup> Adduction Ions**

For the same ten most abundant N-glycan compositions of chicken ovalbumin, MS/MS revealed a total of 18 isomers, IMS-MS revealed a total of 20 isomers and MS<sup>n</sup> revealed a total of 36 isomers. Clearly, MS<sup>n</sup> is more capable of finding glycan isomers. However, for H3N3 composition, 3 isomers were detected by MS<sup>n</sup>, compared with 4 isomers reported by IMS-MS. It is possible that the non-core HexNAc might have different linkages to the core part. During the MS<sup>n</sup> fragmentation process, the non core HexNAc residues are very easy to break down at the very early fragmentation stages. This presents a challenge to find out the linkage of N to the neighboring residue.

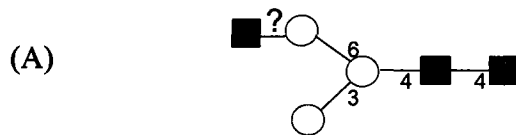
To solve this problem, a special technique was explored in the structure study, e.g., changing adduction ions to fit special requirements. It is known that fragmentation yields are inversely related to the size of adduction cations<sup>35</sup>. An

exploration of adduction ion  $\text{Li}^+$  in H3N3 was used to find out the linkage of non-core HexNAc. For H3N3, most of the terminal HexNAcs lost in the step of  $\text{MS}^3$ . To detect the HexNAc linkages, it is necessary to isolate and disassemble the N attached monosaccharide. Figure 2.19 compares  $\text{MS}^6$ - $\text{MS}^8$  spectra of H3N3 obtained with  $\text{Na}^+$  and  $\text{Li}^+$  adduction ions, step by step. The spectra showed that  $\text{Na}^+$  ion is too “big” to be adducted by small monosaccharide. As a result, the spectrum of  $m/z$  227 (pathway:  $1432 \rightarrow 1139 \rightarrow 880 \rightarrow 662 \rightarrow 417 \rightarrow 343 \rightarrow 227$ ) is not detectable. However, when  $\text{Li}^+$  ions were used, the spectrum of  $m/z$  211 is very informative. It generates the first two structures of H3N3.

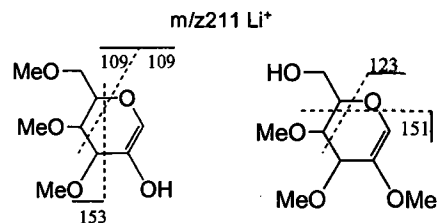
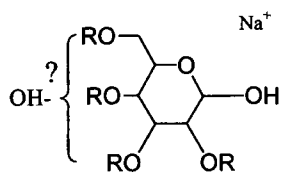
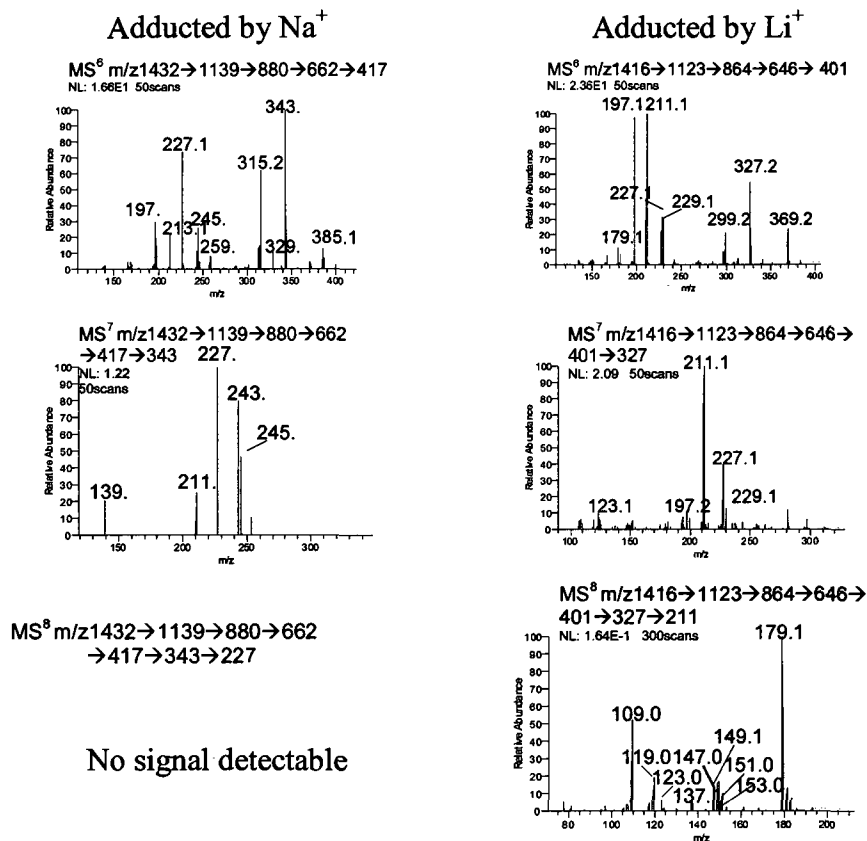
By analyzing mono-residue  $m/z$  211 ( $\text{B}_1$ -ion with  $\text{Li}^+$ )  $\text{MS}^8$  spectrum (pathway:  $1416 \rightarrow 1123 \rightarrow 864 \rightarrow 646 \rightarrow 401 \rightarrow 327 \rightarrow 211$ ), the fingerprint ion  $m/z$  151 indicates a hydroxyl (a scar formed by the loss of HexNAc earlier) on the 6<sup>th</sup> position. This is the diagnostic peak for structure 1 shown in Figure 2.19. Meanwhile, the second dominant ion with  $m/z$  109 discloses the fragments which might come from a hydroxyl at the 2<sup>nd</sup> or 3<sup>rd</sup> position. The ion of  $m/z$  153 formed by a neutral loss of  $m/z$  58 is relatively low in abundance (characterized by 2 position hydroxyl). This indicates the presence of 1-2 structural linkage.

This experiment shows that different metal adduct ions can be used under different circumstances. For small oligosaccharides structures determination, a buildup of standard ( $\text{Na}^+$  and  $\text{Li}^+$ ) library data will be very beneficial.





(B)



Structure 1

Structure 2

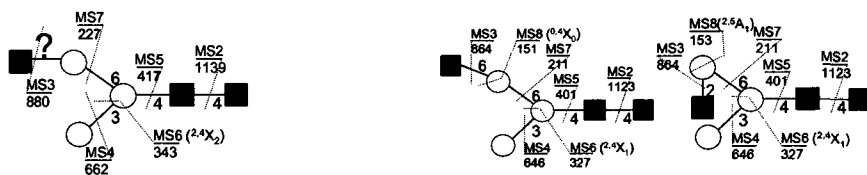


Figure 2.16: (A) An ambiguous topology of H3N3. (B) A step by step comparison of MS<sup>6</sup> – MS<sup>8</sup> spectra obtained with Na<sup>+</sup> and Li<sup>+</sup> adduction ions.

## **2.4 Conclusion**

The use of ion trap mass spectrometry precursor isolation coupling with sequential disassembly to determine the carbohydrate isomeric structures has been described. The presence of carbohydrate isomers has been less appreciated and frequently overlooked. The underlying structural isomers have highly similar structures with very similar physical and chemical properties. This makes it extremely difficult, if not impossible at all, to separate carbohydrate isomers using any types of chromatography. Additionally, because of the inherited limits, any other MS based methods without higher degree fragmentation are blind to the presence of many isomers, which are quite easy to be detected with our approach. This is clearly indicated by the comparison result of our findings and the previously reported structures. Our results strongly suggest that the preeminence of ion trap  $MS^n$ , coupled with the appropriate sample preparation methods and the sophisticated data acquisition/analysis techniques, makes it an ideal instrument for carbohydrate sequencing with minimum assumptions on the underlying structures required.

## CHAPTER III

# GLYCAN BIOMARKER DISCOVERY OF ADULT STEM CELLS USING SEQUENTIAL MASS SPECTROMETRY

### 3.1 Introduction

#### Stem Cell, Adult Stem Cell, and Asymmetric Self-renewal

Stem cells are unspecialized cells which can self-renew by cell division for a long time. Under selected environment, stem cells can be induced to form cells with specific functions, for example, develop into different types of cells, such as muscle, blood, and nerve cells. Such cells can be used to repair damaged tissues. Scientists have been trying to find a way to transplant stem cells to treat diseases and disorders<sup>36-38</sup>.

In general, there are two types of stem cells: embryonic stem cells (ESCs) and adult stem cells (ASCs). ESCs are usually extracted from embryos with the potential of unlimited expansion and pluripotency (ability to differentiate into nearly all types of body cells). Ethical concerns and transplant rejection of ESCs are parts of the hurdles for researchers to deal with. ESCs have been proposed as promising candidates for future therapies. Nevertheless, after nearly ten years of research, there are still no approved treatments using embryonic stem cells<sup>39</sup>.

On the other hand, adult stem cells, also called somatic stem cells, are found in differentiated cells in many organs and tissues. ASCs can differentiate to yield many types of specialized cells without ethical controversies and lower risk of transplant rejection. ASCs from different sources, including umbilical cord blood and bone marrow, have been clinically used in medical treatment for years<sup>41, 42</sup>.

Asymmetric self-renewal is the defining property of ASCs and it is essential for the maintenance of adult stem cell<sup>43</sup>. Following asymmetric cell self-renewal process, ASCs divide to yield one daughter cell, which is a long-lived cell identical to its parent cell, and another daughter cell called progenitor cell. The progenitor cell loses pluripotency and the asymmetric self-renewal attribute. Its further divide to generate specialized cells.

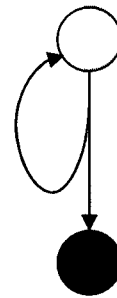


Figure 3.1 Asymmetric ASCs self-renewal

○ -stem cells  
● -progenitor cell

### **Poor History of ASCs Biomarkers Discovery**

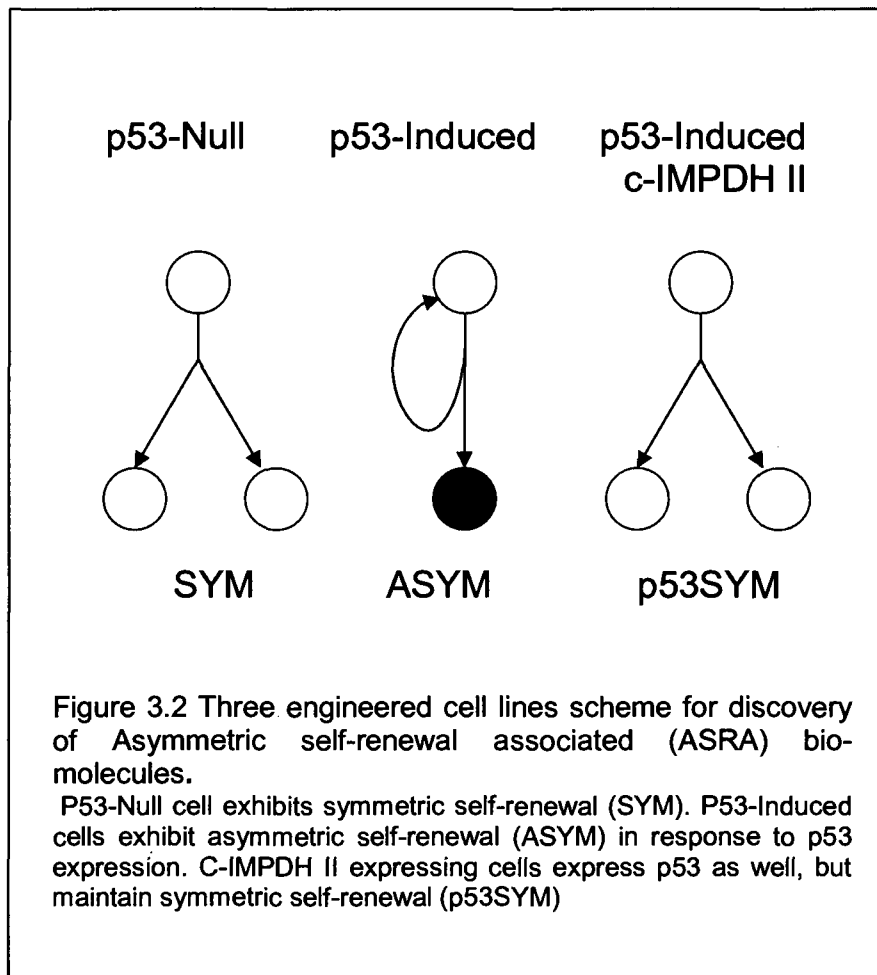
ASC-specific biomarkers are invaluable in many research fields. Stem cell biologists need them to study ASC functions; cancer biologists want a way to evaluate the roles of ASCs as carcinogenic targets; pharmaceutical clinicians desire diagnostic and prognostic markers and indicators of drug-tissue interactions.

However, to date, no protein markers have been identified as tissue-specific for ASC; and also no universal biomarkers have been established to differentiate ASCs from other tissues as well<sup>43</sup>. The difficulty has been due to a number of factors: Firstly, ASCs constitute only a small fraction of tissue cells, estimated to be less than one per several thousands in many tissues, even as low as 1/100,000<sup>44</sup>. Additionally, in culture, cell populations enriched for ASCs decline in purity rapidly. Finally, the defining morphological attributes of ASCs have remained undescribed.

### **A New System for Studying ASCs**

In order to overcome the obstacles outlined above, the Sherley lab at Boston Biomedical Research Institution (BBRI) took a novel approach, which is based on *the hypothesis that some biomarkers specific to asymmetric self-renewal will also be specific to ASCs*. A system that consists of cell lines engineered to recapitulate asymmetric self-renewal under experimental control<sup>45-47</sup>, has been established.

These cell lines are built from p53-null murine fetal fibroblasts. They are engineered to express normal levels of the wild-type p53 gene from a modified metallothionein promoter. In maintenance culture medium (i.e. Zn-free), the conditional p53 gene is off, and the cells undergo symmetric self-renewal. In ZnCl<sub>2</sub>-supplemented medium, normal levels of p53 protein are produced, and the cells switch to asymmetric self-renewal (Figure 3.2, p53-Induced). In recent



gene expression profiling studies, the Sherley lab have shown that the asymmetric self-renewal state of the engineered cell lines has gene expression profiles similar to freshly isolated human ASCs (2009, Sherley, et al, unpublished data). The type II inosine-5'-monophosphate dehydrogenase (IMPDH II), a p53-dependent down-regulation enzyme, was discovered to be necessary for asymmetric self-renewal by engineered cell lines. A congenic derivatives of the original p-53-inducible cell lines constitutively expressed an IMPDH II transgenic cDNA ("c-IMPDH")<sup>48-53</sup>. The c-IMPDH II lines retain Zn-inducible p53 expression,

but do not shift from symmetric self-renewal to asymmetric self-renewal ( scheme 3.2. p53-Induced c-IMPDPH II) <sup>46</sup>. Control p53-null cell lines were co-derived by stably transfecting the metallothionein promoter vector without a p53 cDNA insert. These lines retain symmetric self-renewal in Zn-supplemented medium (Scheme 3.2, p53-Null).

In summary, the system includes three cell lines: p53-Null (SYM), p53-Induced (ASYM), and p53-Induced c-IMPDPH II (p53SYM). Asymmetric self-renewal associated biomolecules (ASRAs) are defined as those which are undetectable or produced at low levels in both SYM states, but show high level production in the ASYM state. The goal of the study is to identify those ASRAs

### **Carbohydrate Biomarker Discovery of Adult Stem Cells**

As one of the three major biomacromolecule classes (protein, carbohydrate, and nucleic acid) in cells, carbohydrate plays critical roles in cellular activities. This naturally makes carbohydrate a good target candidate for biomarker discovery<sup>54-57</sup>. However, due to its structural complexity and the lack of efficient analytical tools, the space of carbohydrate biomarkers has remained unexplored. Over the past several years, an innovative integrated analytical/computational approach has been developed in the Reinhold lab at University of New Hampshire. This chapter summarizes the progress in exploring the glycoproteome for ASC-specific biomarkers using sequential mass spectrometry (MSn) technology.

## **3.2 Experimental Methods**

### **Sample preparation**

Three purified proteins (including glycoproteins) samples from three cell lines (SYM, ASYM, p53SYM) were provided by Dr. Sherley's Lab, Boston Biomedical Research Institute (BBRI). Previous samples treatment described as Dr. Sherley's Lab:

Buffer used:

100mM Ammonium Bicarbonate

10mM EDTA

No protease inhibitors were used

Cell Extract:

Trypsinize the cells and pellet the cells (neutralized the trypsin with media and washed twice with PBS)

Add the above buffer

Sonicate the cells and spin @ 4C 13,000rpm for 20min

Collect the supernatant

Protein by Bio-Rad DC protein Assay

The samples were processed under identical conditions in parallel in our lab<sup>10</sup>:

(1) N-linked glycans were released enzymatically by PNGase

(2) N-linked glycans were collected by groups of neutral glycans and acidic glycans.

(3) Glycans were reduced with a solution of sodium borohydride (NaBH<sub>4</sub>).



(4) Permethylation was carried out as described by Ciucanu and Kerek<sup>24</sup>.

(5) Reduced and permethylated glycans were re-suspended in a 75% (v/v) methanol aqueous solution for mass spectrometry analysis.

### **Mass spectrometry**

The ion trap mass spectrometer instrument (LTQ, Thermo Fisher Scientific, Waltham, MA) was equipped with a TriVersa Nanomate® for automatic nanoelectrospray injection. Spectra were collected using Xcalibur 2.0 software (Thermo Fisher Scientific). Collision parameters were left at default values with normalized collision energy set at 35%. Activation Q was set at 0.25, and activation time was set at 30 ms.

### **Data analysis**

In-house developed software was used to assist data analysis. A list of putative glycan ions and corresponding compositions were determined from MS profiles. An ion candidate pool was constructed and ranked based on the statistical significance of intensity fold-changes between samples and the signal intensity values. The top ~50 ions were selected from MS profiles for further MS<sup>n</sup> experiments. A total of ~1000 spectra were collected and analyzed in this study. The spectra with same MS<sup>n</sup> pathways were compared pair-wise among samples. The spectra-pairs with significant difference across samples were selected for further manual validation. The putative structures were then proposed based on glycan fragment compositions.

### **3.3 Results and Discussion**

A systematic differential N-glycome analysis was performed with the three cell lines using the sequential mass spectrometry techniques. For clarity, the results are organized into three subsections: (1) N-glycan MS profile comparison; (2) the discovery of the sample-specific glycan isomers; and, (3) the exploration of N-glycan branching patterns.

#### **3.3.1 N-glycan MS profile comparison**

The ESI MS profiles of N-glycans from the three samples are shown in Figure 3.3. These three profiles have great similarities. However, the profiles still exhibit significant differences between three samples. The distribution of major N-glycans detected from MS profiles is summarized in Table 3.1. The relative abundances of N-glycans were calculated based on the most abundant peak  $m/z$  1115.9 (H8N2) from each profile. The base peak  $m/z$  1115.9 is assigned with a relative abundance value of 100. The glycans with significant fluctuations in relative abundance across cell lines are highlighted by shaded rows.

According to the synthetic process of N-linked glycans, the glycan compositions are divided into four groups<sup>58-61</sup>. H10N2 to H7N2 (early processing stage); H6N2 to H2N2 (paucimannose forming stage); HxNy( $y \geq 3$ ) (hybrid and complex forming stage); HxNyF1 (fucosylation stage). Figure 3.4 is a bar diagram of the relative abundance distribution of HxN2 ( $x \geq 7$ ), HxN2( $x < 7$ ), HxNy( $y \geq 3$ ) and HxNyF1.

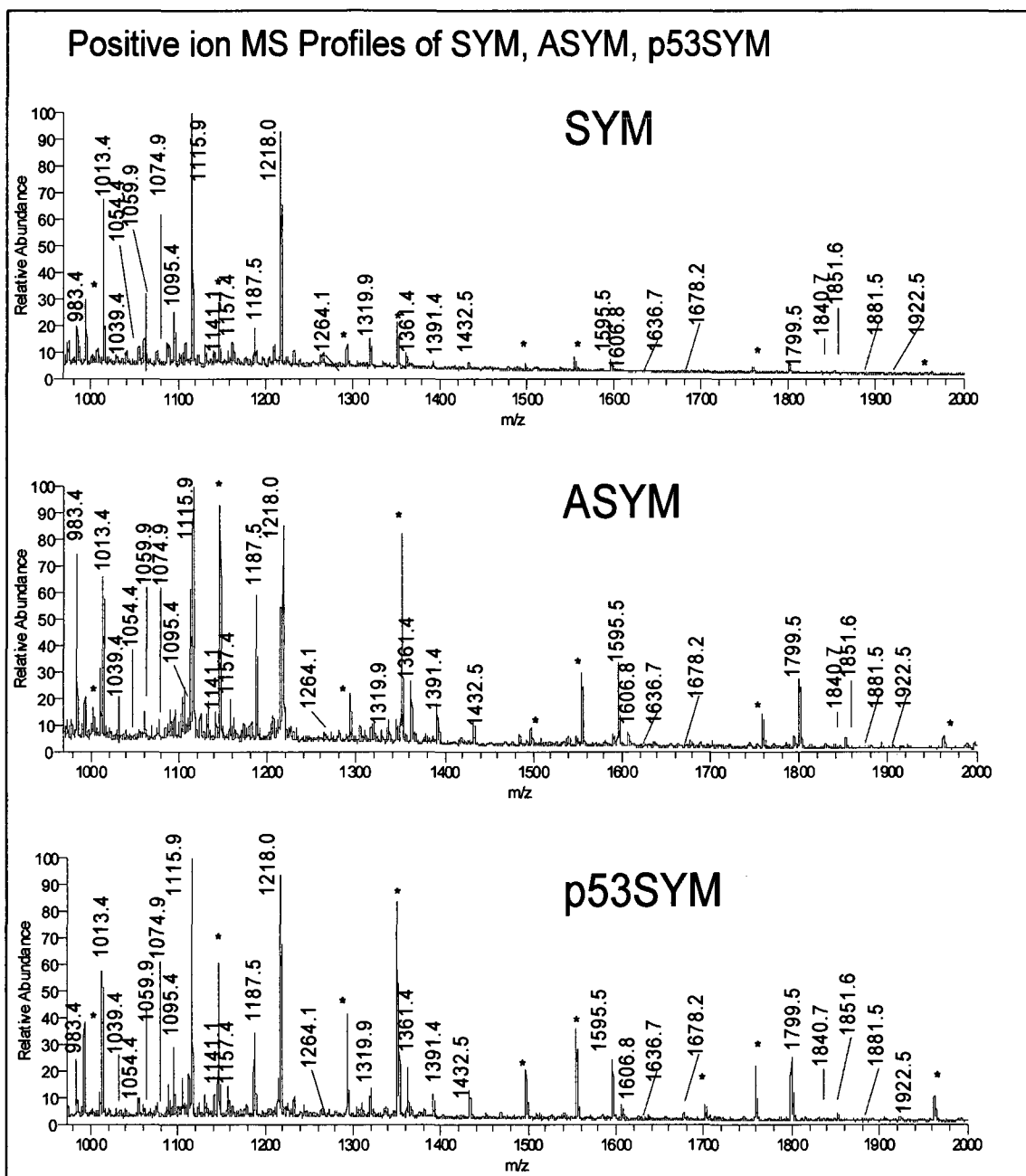


Figure 3.3: Comparative ESI-MS profiles analysis of reduced and methylated neutral N-glycans from three engineered stem cell lines: SYM, ASYM and p53SYM. The peaks labeled with numbers are corresponding to N-glycans, Peaks labeled with a \* are imp peaks (free glycans).

Glycan Group	m/z (Na <sup>+</sup> )	Composition	Relative Abundance		
			SYM	ASYM	p53SYM
Group1: HxN2 (2≤x<7)	983(+)	H2N2	20.5	71.3	33
	1187(+)	H3N2	20.4	57.4	34.2
	1391(+)	H4N2	6.9	17.7	11.5
	1595(+)	H5N2	7.7	32.9	16
	1799(+)	H6N2	5.2	25.6	25.3
Group2: HxN2 (x≥7)	1013(2+)	H7N2	67.8	71.6	57.3
	1115(2+)	H8N2	100	100	100
	1218(2+)	H9N2	92.6	86.5	93.6
	1319(2+)	H10N2	15.4	12.7	12.9
Group3: HxNy (y≥3)	1432(+)	H3N3	6.5	11.5	8.4
	1636(+)	H4N3	1.5	3.6	3.6
	1840(+)	H5N3	2.6	2.8	2.5
	1678(+)	H3N4	1.5	5.5	4.5
	1881(+)	H4N4	2.6	2.6	2.3
	1054(2+)	H5N4	12.6	9.3	10.3
	1922(+)	H3N5	2	3.2	2.8
	1074(2+)	H4N5	10	9.1	7
	1176(2+)	H5N5	8	9	6.1
Group4: HxNyF1	1157(+)	H2N2F1	10.8	19.7	16.2
	1361(+)	H3N2F1	10.3	25.5	21.4
	1606(+)	H3N3F1	4.1	8.1	7.7
	1810(+)	H4N3F1	2.6	2.3	2.8
	1851(+)	H3N4F1	3.2	5	5
	1039(2+)	H4N4F1	11.2	9.3	6.2
	1141(2+)	H5N4F1	10.4	15	11.1
	1243(2+)	H6N4F1	6.4	5.8	6.6
	1161(2+)	H4N5F1	12.6	12.9	6.5
	1264(2+)	H5N5F1	9.5	7.7	6.4
Total	Total abundance		464.9	643.6	521.2

Table 3.1: The distribution of major N-glycans detected from MS profiles. The glycan compositions are grouped according to the synthetic process of N-linked glycans. The glycans with significant fluctuations in relative abundance across cell lines are highlighted by shaded rows.

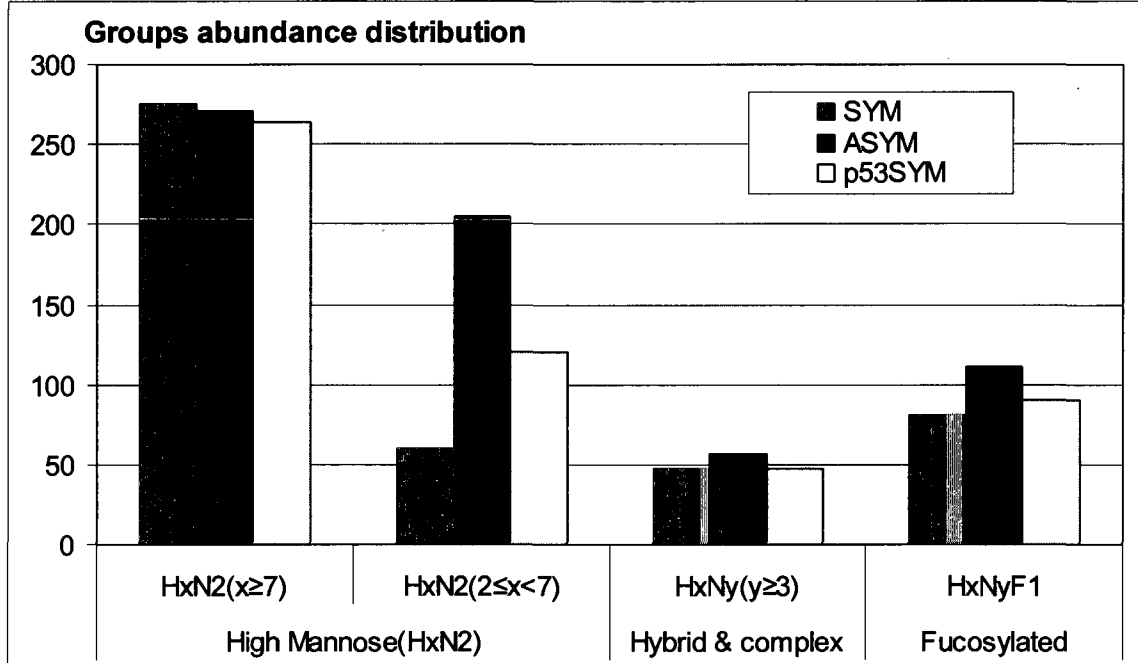


Figure 3.4: The group abundance distribution of N-glycans across the three cell lines: SYM, ASYM, and p53SYM.

For composition groups HxN2 ( $x \geq 7$ ), HxNy ( $y \geq 3$ ) and HxNyF1, it appears that there are no significant differences in these three samples. For example, in HxN2 ( $x \geq 7$ ), the relative abundance is 275.8, 270.8 and 263.8 in SYM, ASYM and p53SYM, respectively. However, for composition group HxN2( $x < 7$ ), the relative abundances are significantly different in SYM, ASYM and p53SYM. In addition, for each composition, such as H2N2 and H3N2, the relative abundances in different samples are significantly different.

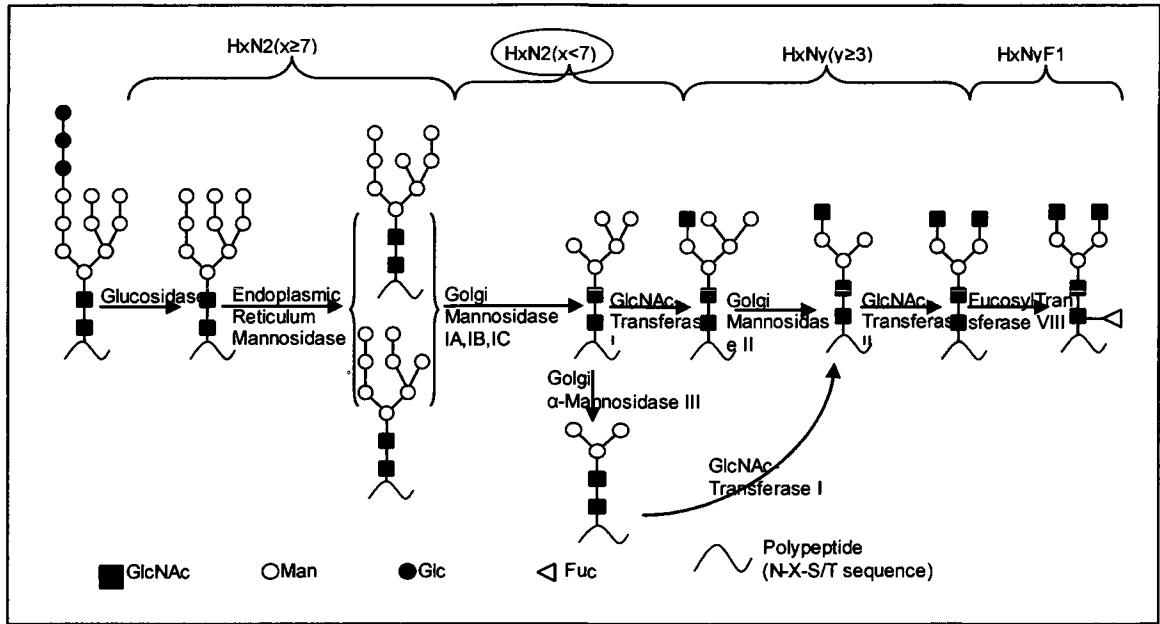


Figure 3.5: The major biosynthesis steps of N-glycans that localized in part of endoplasmic reticulum(ER) and Golgi apparatus of mammalian cells. The categories of N-glycans were divided into HxN2 ( $x \geq 7$ ), HxN2 ( $x < 7$ ), HxNy ( $y \geq 3$ ), HxNyF1, according to the need of the data treatment. Modified from Herscovics A. 1999<sup>58</sup>

The comparison results of MS profiles clearly indicate that the N-glycan biosynthesis activities processed differently in the three cell lines. Figure 3.5 shows the major biosynthesis steps of N-glycans that are localized in part of the endoplasmic reticulum (ER) and Golgi apparatus of mammalian cells. Various enzymes participate in each N-glycan synthetic step<sup>58</sup>. Among them, Golgi  $\alpha$ -Mannosidase IA, IB, IC and Golgi  $\alpha$ -Mannosidase III are enzymes closely related to the differences of group of HxN2 ( $x < 7$ ) N-glycans based on profile comparison. Our results suggest that the enzyme activities and/or the gene expression levels of these related enzymes should be different in the three cell lines. This prediction can be verified by enzymatic activity assays and microarray experiments.

### **3.3.2 The discovery of the sample-specific glycan isomers**

Most of the around 1000 MS<sup>n</sup> spectra collected showed great similarities between samples at every MS<sup>n</sup> level. This suggested that the three samples were highly correlated and the underlying structures were considerable similar between samples. However, isomeric structural differences have been detected. Here, the spectra and the structures identified within the ion m/z 1810.6<sup>+</sup> served as an example for illustration purpose.

In this study, the ion m/z 1810.6<sup>+</sup> which corresponds to N-linked glycan composition H4N3F1, was selected as an example to illustrate the structural identification and differentiation process.

The ion m/z 1810.6<sup>+</sup> (H4N3F1) was isolated from each sample and processed on LTQ-MS. Table 3.2 shows the side by side comparison of MS<sup>n</sup> spectra of the samples. The ions that form diagnostic pathways were underlined in the spectra. The MS<sup>n</sup> spectra remain almost identical among samples until the MS<sup>4</sup> / MS<sup>5</sup> level. Table 3.3 summarizes five isomers identified from the ion m/z 1810.6<sup>+</sup> (H4N3F1) and their distributions within these three samples. The diagnostic pathways and corresponding glycan fragments are also shown in this table.

#### **Common structures shared by all three samples**

The structure S1 is defined by the pathway 1810 → 1343 → 880 → 662 → 458 and 1810 → 1343 → 486 → 268 (259). Although the relative intensities

might be different, this pathway can be detected in all three samples. The structure S1 is the major glycan compound in all samples.

The ion of  $m/z$  1343 in  $MS^2$  spectrum corresponds to a neutral loss of reducing N with a terminal F attached. In  $MS^3$ , the complement ions  $m/z$  486 and 880 were observed. The ion of  $m/z$  486 suggests that one N and one H should bundle together as a terminal. However, it is not clear whether N or H is the terminal residue at this level. Further fragmentation of the ion  $m/z$  486,  $MS^4$  (the notation  $MS^4$  is used here in order to differentiate it from another  $MS^4$  of H4N3F1 fragments) generated the complement ions  $m/z$  268 and  $m/z$  259, which indicate terminal H and internal N, respectively. The ion  $m/z$  880 (complementary to  $m/z$  486) was disassembled; the daughter ion  $m/z$  662 in  $MS^4$  is the result of losing a terminal H. With the N-linked core structure constraint, the ion  $m/z$  458 in  $MS^5$  (neutral loss of internal H) finally defined the structure S1 in Table 3.3.

Another common structure shared by the samples is determined as structure S2 (shown in Table 3.3). Its diagnostic pathway is  $1810 \rightarrow 1343 \rightarrow 1084 \rightarrow 880 \rightarrow 445$ . The neutral losses are: reducing N + terminal F (one scar), terminal N (one scar), internal H (two scars) and H+N with three scars, respectively. The leftover  $m/z$  445 is terminal H2 dimer. With the constraint of the N-glycan core structure, the structure S2 can be unambiguously determined.

### **Structure shared by SYM and ASYM**

The pathway  $1810 \rightarrow 1343 \rightarrow 1084 \rightarrow 880 \rightarrow 676$  can only be found in SYM and ASYM samples. The sequence of neutral losses corresponds to a loss



of a reducing end N +terminal F (one scar), terminal N, internal H and another internal H, respectively. This indicates the sugar sequence of a branching arm should be N-H-H-. With the definition of N-glycan core structure, the structure S3 is determined (as shown in Table 3.3). According to literature<sup>2, 58, and 59</sup>, this is not a usual structure corresponding to the major N-glycan biosynthesis.

### **Structure shared by ASYM and p53SYM**

Similarly, the pathway 1810 → 1343 → 1084 → 866 → 676 can be observed in the ASYM and p53SYM samples, but is not present in the SYM sample. The ion m/z 1084 in MS<sup>3</sup> is the result of losing a terminal N; In MS<sup>4</sup>, the ion m/z 866 indicates the presence of another terminal H. The ion m/z 676 in MS<sup>5</sup> is a loss of one H with three scars, which must be connected with the early neutral loss of terminal N and H. The structure S4 can be defined by this diagnostic MS<sup>n</sup> pathway.

### **ASYM specific isomer**

In the MS<sup>5</sup> spectrum of the pathway 1810 → 1343 → 1084 → 866, the ion m/z 444 and m/z 463 are found in the ASYM sample only, but undetectable in the SYM and p53SYM samples, as shown in Table 3.2. The ion m/z 1084 in MS<sup>3</sup> and

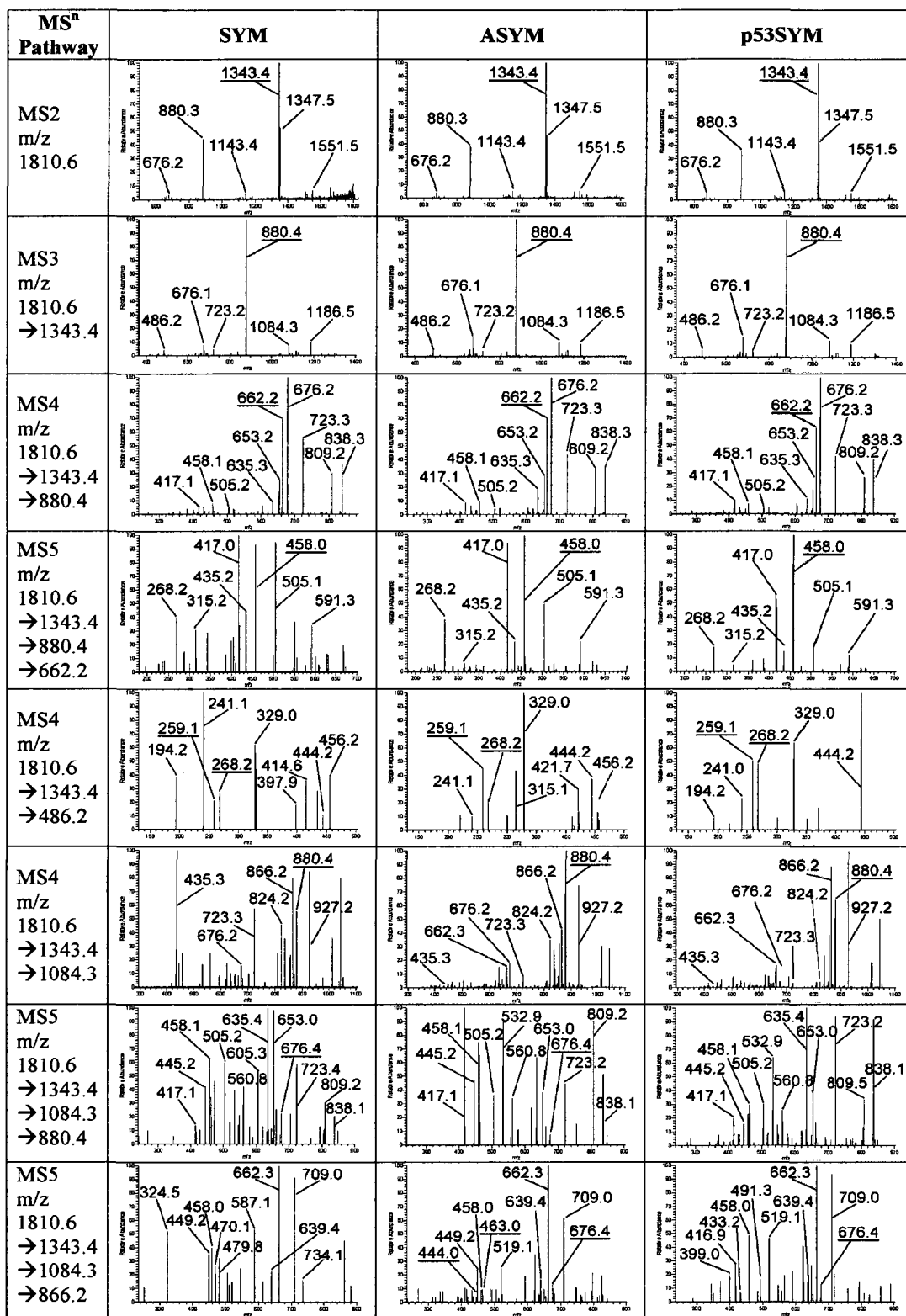


Table 3.2 Side by side comparison of MS<sup>n</sup> spectra obtained from three cell lines within the ion m/z 1810.6+ (H4N3F1). The ions corresponding to the diagnostic pathways are highlighted in the spectra.

ID	Glycan structure and diagnostic pathway	SYM	ASYM	p53SYM
S1	<p><b>MS<sup>n</sup> Pathways:</b>  1810 → 1343 → 880 → 662 → 458  1810 → 1343 → 486 → 259 / 268</p>	++	++	++
S2	<p><b>MS<sup>n</sup> Pathway:</b>  1810 → 1343 → 1084 → 880 → 445</p>	+	+	+
S3	<p><b>MS<sup>n</sup> Pathway:</b>  1810 → 1343 → 1084 → 880 → 676</p>	+	+	-
S4	<p><b>MS<sup>n</sup> Pathway:</b>  1810 → 1343 → 1084 → 866 → 676</p>	-	+	+
S5	<p><b>MS<sup>n</sup> Pathway:</b>  1810 → 1343 → 1084 → 866 → 463 / 444</p>	-	+	-

Table 3.3: The distribution of the isomers identified within the ion  $m/z$  1810.6+ (H4N3F1). The diagnostic pathways and the corresponding glycan fragments are labeled in the diagram. (++: present as a major compound; +: present; -: undetected)

the ion  $m/z$  866 in  $MS^4$  indicate there should be a terminal N and a terminal H lost. The ions  $m/z$  444 and 463 in  $MS^5$  are complement fragments. This suggests the presence of another terminal dimmer H2. By using the N-linked core constraint, and putting all the residues together, the structure S5 can be derived with minimal ambiguity. This ASYM-specific glycan structure is qualified to be considered as a candidate of the asymmetric self-renewal associated biomolecules (ASRAs).

### **3.3.3 The exploration of N-glycan branching patterns**

By investigation of a large N-glycans pool, not only ASYM specific N-glycans were revealed, but also different kinds of structural abundance trends were discovered among three different cell lines.

Most of the spectra corresponding to each composition from different samples display a very similar cleavage pattern. However, the quantitative differences between each sample are still present. This quantitative difference appears trivial and useless when analyzed separately. However, after data collection and re-arrangement, the trend of certain glycosylation patterns was discovered. The ability of LTQ instrument quantitative definition has not been qualified. However, this relative quantity comparison might lead to the discovery of glycosylation mechanisms or enzyme functions.

## Structure templates

To understand the pattern of structures in each category, three templates were set up based on the structures that we investigated. Figure 3.6 shows the proposed structural templates. They do not include any specific linkages. These templates may reflect different functions of various GlcNAc-transferases among samples.

Template 1 represents the type of N-glycan structures with following features: one N was attached to the core 1-3 (or 1-6) arm Mannose and no N was attached to the inner core Mannose; the other arm could be N or NH dimer attached or no other residues attached; the reduced terminal N could be fucosylated or not. The core fragments from stage  $MS^n$  spectra with  $m/z$  662 and  $MS^{n+1}$  spectra with  $m/z$  458 reflect the abundance of structures of this category.

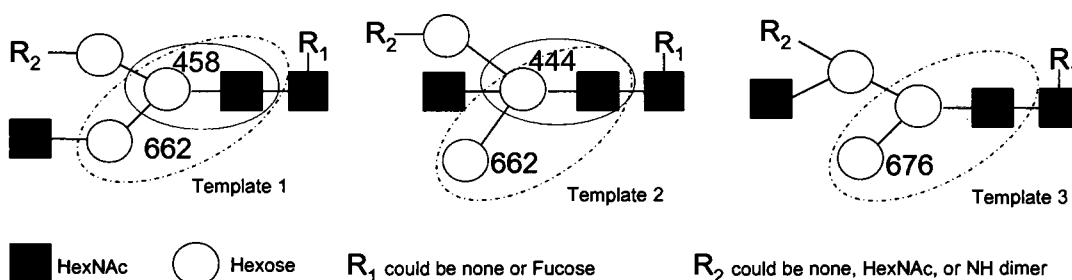


Figure 3.6 Structural templates. The linkages were not specified.

Template 2 represents the type of N-glycan structures with following features: one N attached to the inner core Mannose; one arm could be attached by N, NH or no attachment and no attachment on the other arm; the reduced

terminal N could be fucosylated or not. The core fragments from stage  $MS^n$  spectra with  $m/z$  662 and  $MS^{n+1}$  spectra with  $m/z$  444 reflect the abundance of structures of this category.

Template 3 represents the type of N-glycan structures with following features: there is a branch on one core arm- one N and another N (or NH dimer) attached to this arm; no attachment to the other arm; the reduced terminal N could be fucosylated or not. The core fragments from certain stage  $MS^n$  spectra with an  $m/z$  value of 676 will reflect the abundance of structures of this category.

### **Structures relative abundance distributions**

A total of 28 N-glycan compositions were investigated for the detailed structures. Based on the proposed templates, data of those fitting structures were collected from specific peaks of certain stage  $MS^n$  spectra to get the relative abundances of each categorical structures. Data obtained from each structure were taken from the corresponding spectra of three samples at the same condition. By analyzing these data, two structural trends were found.

Trend 1: Table 3.4 shows relative isomeric abundances of 5 compositions collected from characteristic signature ions. For example,  $m/z$  458 was selected for Template 1, and  $m/z$  444 was selected for the structure of Template 2. The corresponding fragmentation pathway for signature ions of each composition is also shown in the table. The measurement of relative abundance of each ion is based on the signature ion  $m/z$  458 of Template 1, i.e., the relative abundance or ratio of the ion  $m/z$  458 is 100%. The last three columns of Table 3.5 are

relative abundances of structures of Template 2 for different sample compositions.


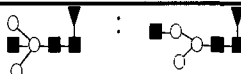


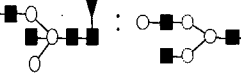
	m/z & composition	pathway & signature ions	Structure Category template2: template1	Ratio of the signature ions		
				SYM	ASYM	p53SYM
1	1432 <sup>+</sup> H3N3	1432->1139->880->662 444:458		9.65%	11.80%	1.98%
2	1606 <sup>+</sup> H3N3F1	1606->1139->880->662 444:458		3.85%	2.70%	0.93%
3	1677 <sup>+</sup> H3N4	1677->1418->1125->866->662 444:458		35.60%	40.00%	18.40%
4	1851 <sup>+</sup> H3N4F1	1851->1384->1125->866->662 444:458		8.90%	12.40%	6.70%
5	1039 <sup>2+</sup> H4N4F1	1039->909->1333->866->662 444:458		21%	23%	2.40%

Table 3.4 Comparison of structural isomers (Template 1 and Template 2) distribution in each composition of samples: SYM, ASYM, p53SYM.

The structural abundance ratios of Template 2 versus Template 1 show a trend among the samples. It is shown that SYM and ASYM cell lines have similar “tendency” to form Template 2 style of structures. The tendency appears much stronger than that of p53SYM. This trend might be an indicator (s) of functions of different cells, and / or the relative strengths of corresponding enzymes. Significant differences of structural distribution are observed in a particular composition. For example, H4N4F1 has a relative abundance ratios of 21%, 23% and 2.4% in SYM, ASMY and P53SYM, respectively. This might be used as a biomarker to identify and isolate cells of p53SYM from other cells. Figure 3.7 shows the structural distributions of three samples for different compositions.

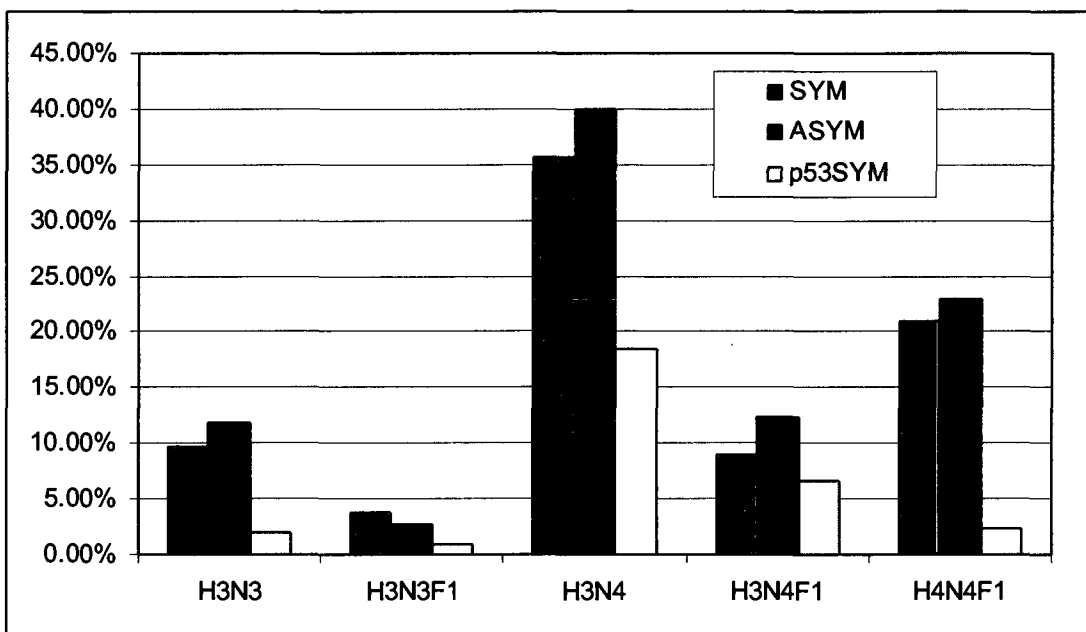


Figure 3.7 Comparison of relative abundance of Template 2 structures versus Template 1 structures among sample SYM, ASYM and p53SYM.

Trend 2: Table 3.5 lists the relative abundances of 5 compositions from Template 3 and Template 1 and 2. The fragmentation pathways for each composition of each ion are shown in Table 3.6. Similarly, the relative abundances were calculated by ratioing the abundance of each composition with the abundance of the characteristic signature ion. In this case, the ion of  $m/z$  662 was selected for the structures of Template 3, and the ion of  $m/z$  of 676 was selected for Template 1 and 2. The last three columns of Table 3.5 are the relative abundances of structures of Template 3 for each composition.



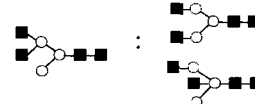
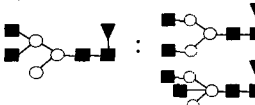
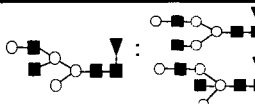
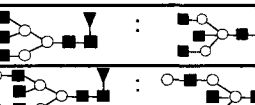
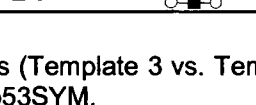
	m/z & composition	pathway & signature peak	Structure Category template3: template1+2	Ratio of the signature ions		
				SYM	ASYM	p53SYM
1	1677 <sup>+</sup> H3N4	1677->1418->1125->866 676:662		3.85%	4.45%	5.30%
2	1851 <sup>+</sup> H3N4F1	1851->1384->1125->866 676:662		1.03%	2.20%	1.93%
3	1039 <sup>2+</sup> H4N4F1	1039->909->1333->866 676:662		0.99%	7.16%	1.41%
4	1059 <sup>2+</sup> H3N5F1	1059->930->800->671->852		10.47%	13.05%	26.00%
5	1263 <sup>2+</sup> H5N5F1	1263->1134->902->1319->852 662:648		2.87%	40.20%	36.30%

Table 3.5 A comparison of structural isomers (Template 3 vs. Template 1+2) distribution in each composition among samples: SYM, ASYM, p53SYM.

The structural abundance ratios of Template 3 to Template 1 and 2 are quite different in different samples. SYM always has a lower abundance of Template 3 structures than ASYM and p53SYM. For some specific compositions, such as H4N4F1, H3N5F1 and H5N5F1, the differences are fairly big. These differences might be used as biomarkers to identify and isolate different stem cells. Figure 3.8 shows the structural distributions of three samples for different compositions.

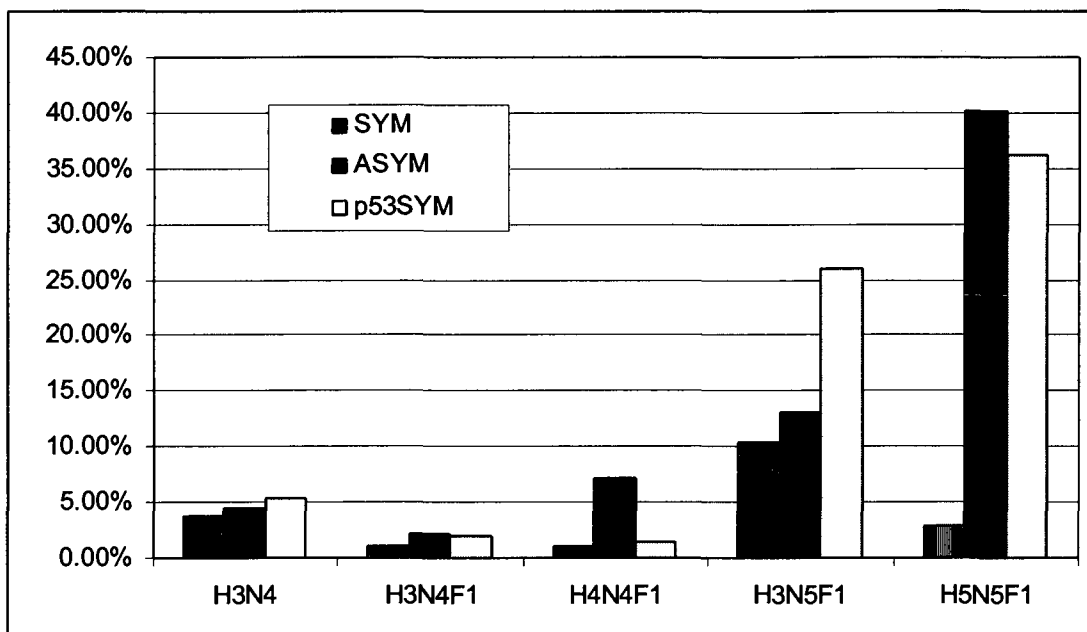


Figure 3.8 Comparison of relative abundance of Template 3 structures versus Template 1 + 2 structures among sample SYM, ASYM and p53SYM.

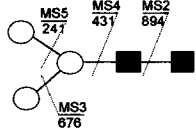
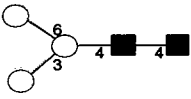
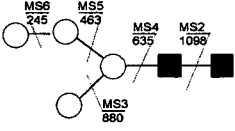
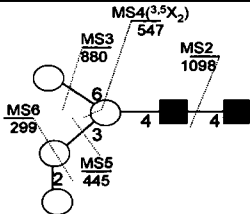
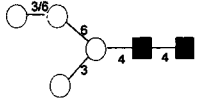
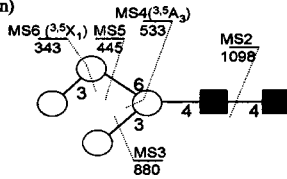
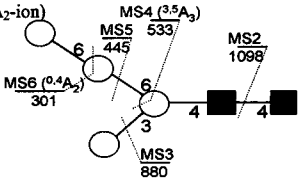
Different GlcNAc transferases have been isolated and characterized in N-glycans branching of animal cells (Figure 3.9). For identical substrates, the glycosyltransferases compete on given N-glycans to make the products varied<sup>2, 60-64</sup>. This might be the reason for different structural distributions of these three samples. GlcNAc T-III appears more competitive in sample SYM and sample ASYM than in p53SYM. Meanwhile, GlcNAc T-VI and GlcNAc T-V exhibits weaker strengths in sample SYM, than in sample ASYM and sample p53SYM. The differences unveiled in this study might provide clues for further biological exploration.



Importantly, these data again demonstrate that isomeric differences are often not exposed until MS<sup>3</sup> through MS<sup>5</sup>. And, as such, these data strongly suggest that routine tandem MS (i.e. MS<sup>2</sup>) in many cases would be incapable of defining these differences. MS<sup>n</sup>, when coupled with appropriate sample preparation, however appears uniquely capable of not only identifying this complexity, but fully characterizing each structural difference in the absence of time consuming chromatographic efforts. This makes MS<sup>n</sup> technology an invaluable tool for comprehensive carbohydrate analysis, HTP analysis and probable automation.

## APPENDIX

### Isomeric structures determination by MS<sup>n</sup> (10 compositions of N-glycans from chicken ovalbumin)

Glycan composition	Isomers by MS <sup>n</sup>		Isomers reported (references)
	Topology Isomers	Linkage Isomers	
H3N2 m/z 1187.5 (Na <sup>+</sup> )	1187→894→676→431→241 		 (reference 15)
H4N2 m/z 1391.6 (Na <sup>+</sup> )	1391→1098→880→635→463→245 	1391→1098→880 →547( <sup>3,5</sup> X <sub>2</sub> -ion) →445→ 299( <sup>2,5</sup> X <sub>1</sub> -ion) 	 (reference 15)
		1391→1098→880→533( <sup>3,5</sup> A <sub>3</sub> -ion)→ 445→343( <sup>3,5</sup> X <sub>1</sub> -ion) 	
		1391→1098→880→533( <sup>3,5</sup> A <sub>3</sub> -ion)→ 445→301( <sup>0,4</sup> A <sub>2</sub> -ion) 	

(To be continued)

Table: Summary information of topology and linkages determination for each composition by MS<sup>n</sup>. Comparison structures reported by documents also was included.

Notes:

1. The topology isomers listed didn't involve linkages information. The linkage isomers gave the full or partial details of linkage information depending on sample abundance, isomer structure and necessity. The linkages were labeled if enough evidences were spotted. The path way used in topology detection may be changed for the linkage determination by demands

2. References structures information involved two papers (reference 28&31). The structures from reference31 may be involved the linkages clarified without notification. The structures from reference28 may be reported by other investigators (mainly NMR evidence) without specification from the authors.

<p>H3N3 m/z 1432.6 Na<sup>+</sup> and m/z 1416.6 Li<sup>+</sup></p>	<p>1432→1139→880→ 676→458→268</p>	<p>(Li+) 1416→1123→864→646→401→ 327(<sup>24</sup>X<sub>1</sub>-ion)→211→ 153(<sup>25</sup>A<sub>1</sub>-ion) 109(<sup>35</sup>A<sub>1</sub>, <sup>35</sup>X<sub>0</sub>-ion),</p>	<p>(reference 31)</p> <p>(reference 31)</p>	
	<p>1432→1139→880→ 662→444→268</p>	<p>(Li+) 1416→1123→864→646→401→ 327(<sup>24</sup>X<sub>1</sub>-ion)→211 →151(<sup>04</sup>X<sub>0</sub>-ion)</p>		<p>(reference 31)</p> <p>(reference 31)</p>
	<p>1432→1139→880→662 →417→329(<sup>35</sup>X<sub>1</sub>-ion) →245</p>	<p>(Na+)1432→1139→880→662 →417→329(<sup>35</sup>X<sub>1</sub>-ion) →245</p>		
<p>1595→1302→1084→839→667 →449→241</p>	<p>1432→1139→880→662 →417→315(<sup>35</sup>A<sub>2</sub>-ion) →259</p>	<p>(reference28&amp;31)</p>		
<p>1595→1302→1084→839→635→431</p>	<p>1432→1139→880→662 →417→259(or241)</p>		<p>(reference28&amp;31)</p>	
<p>1595→1302→1084→839→635→ 463</p>	<p>1595→1302→1084→839→737(<sup>35</sup>A<sub>3</sub>-ion)→667 →449→259(or241)</p> <p>(also referred m/z 667 and m/z 449 library data)</p>			<p>(reference28&amp;31)</p>
<p>1595→1302→1084→839→635→ 463</p>	<p>1595→1302→1084→839→737(<sup>35</sup>A<sub>3</sub>-ion)→667 →449→227</p>	<p>(reference28&amp;31)</p>		
<p>1595→1302→1084→839→635→ 463</p>	<p>1595→1302→1084→839→635→533(<sup>35</sup>A<sub>3</sub>-ion) →463</p>		<p>(reference28&amp;31)</p>	

(To be continued)


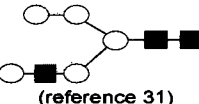
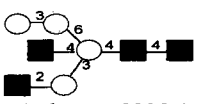
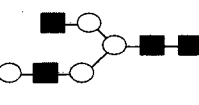
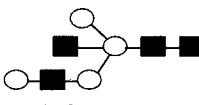
<p>H4N3 m/z 1636.6 (Na<sup>+</sup>)</p>	<p>1636 → 1343 → 1084 → 866 → 621 → 463</p>	<p>1636 → 1343 → 1084 → 866 → 621 → 519 (<sup>3,5</sup>A<sub>3</sub>-ion) → 463</p>	<p>(reference 31)</p>
	<p>1636 → 1343 → 1084 → 866 → 621 → 547 (<sup>3,5</sup>X<sub>2</sub>-ion) → 463</p>	<p>1636 → 1343 → 1084 → 866 → 621 → 547 (<sup>3,5</sup>X<sub>2</sub>-ion) → 463</p>	
<p>H4N3 m/z 1636.6 (Na<sup>+</sup>)</p>	<p>1636 → 1343 → 1084 → 880 → 635 → 463</p>	<p>1636 → 1343 → 1084 → 880 → 635 → 533 (<sup>3,5</sup>A<sub>3</sub>-ion) → 463</p>	<p>(reference 28)</p>
	<p>1636 → 1343 → 1084 → 880 → 635 → 547 (<sup>3,5</sup>X<sub>2</sub>-ion) → 463</p>	<p>1636 → 1343 → 1084 → 880 → 635 → 547 (<sup>3,5</sup>X<sub>2</sub>-ion) → 463</p>	
	<p>1636 → 1343 → 1125 → 880 → 676 431 → 241</p>	<p>1636 → 1343 → 1125 → 880 → 676 → 431 → 241 and 329 (<sup>3,5</sup>A<sub>2</sub>-ion)</p>	<p>(reference 31)</p>
	<p>1636 → 1343 → 1084 → 880 → 635 → 431</p>		

(To be continued)

<p><b>H3N4</b> m/z 1677.6 (Na<sup>+</sup>)</p>	<p>1677→1384→1125 →866→662→458 →268</p>	<p>1677→1384→1125→866→662→417 →315(<sup>3,5</sup>A<sub>2</sub>-ion)→245</p>	<p>(reference 31)</p>
	<p>1677→1384→1125 →866→648 →444→268</p>	<p>1677→1384→1125→866→662→417 →315(<sup>3,5</sup>A<sub>2</sub>-ion)→259</p>	<p>(reference 28&amp;31)</p>
	<p>1677→1384→1125 →866→648→458 →268</p>	<p>1677→1384→1125→866→662→417 →343(<sup>3,5</sup>X<sub>1</sub>-ion)→241</p>	<p>(reference 28)</p>
	<p>1677→1384→1125 →866→648→458 →268</p>	<p>1677→1384→1125→866→676→431 →329(<sup>3,5</sup>A<sub>2</sub>-ion)(and 213)</p>	
<p><b>H6N2</b> m/z 1799.7 (Na<sup>+</sup>)</p>	<p>1799→1506→1084→839→667→449 →259</p>	<p>1799→1506→1084→839→737(<sup>3,5</sup>A<sub>2</sub>-ion)→667 →449→259 (667 and 449 referred library data)</p>	<p>(reference 28&amp;31)</p>
	<p>1799→1506→1084→839→667→449 →227</p>	<p>1799→1506→1084→839→737(<sup>3,5</sup>A<sub>4</sub>-ion)→667 →449→227</p>	
	<p>1799→1506→1084→839→635→ 431</p>		
	<p>1799→1506→1084→866→676</p>		

(To be continued)



H5N3 m/z 1840.8 (Na <sup>+</sup> )	1840→1547→1288 1070→852→634→ 444→268	1840→1547→ 1288→1070→825→723( <sup>3,5</sup> A <sub>3</sub> )→667 667→449 (library data matching)	
	1840→1547→ 921→662→458→268 667→449→259	1840→1547→ 1288→1084→839→737( <sup>3,5</sup> A <sub>3</sub> )→667 667→449 (library data matching)	 (reference 28&31)
H5N3 m/z 1840.8 (Na <sup>+</sup> )	1840→1547→ 1084→880→635→463 486→268		 (reference 31)
	1840→1547→ 880→635→431 690→463		
H4N4 m/z 1881.8 (Na <sup>+</sup> )	1881→1588→1329 →1070→866→ →621→463	1881→1588→1329→1070→866→621 →519( <sup>3,5</sup> A <sub>3</sub> -ion)→463	 (reference 28&31)
	1881→1588→1329 →1070→880→ 635→463	1881→1588→1329→ 1070→866→621 →547( <sup>3,5</sup> X <sub>2</sub> -ion)→463	
	1881→1588→1329 →1070→880→ 635→463		 (reference 31)
	1881→1588→ 921→662→417→227 690→486→268		 (reference 31)

(To be continued)

<p>H3N5 m/z 1922.9 (Na<sup>+</sup>)</p>	<p>1922→1663→1370→1111→852→ 662→444→268</p>	<p>(reference28&amp;31)</p>
	<p>1922→1663→1370→1111→852→ 648→444→268</p>	<p>(reference28&amp;31)</p>
	<p>1922→1663→1370→1111→852→ 662→458</p>	

(Summary Table End)

## REFERENCES

1. Nelson, David L.; Cox, Michael M. (2005). *Lehninger Principles of Biochemistry*, 4<sup>th</sup> edition.
2. Varki, A.; Cummings, R.D.; Esko, J.D.; Freeze, H.H.; Stanley, P.; Bertozzi, C.R.; Hart, G.W.; Etzler, M.E.,(2008) *Essentials of Glycobiology* 2<sup>nd</sup> ed.
3. Varki A. Biological roles of oligosaccharides: All of the theories are correct. *Glycobiology*. 1993; 3: 97–130
4. Apweiler R, Hermjakob H, Sharon N. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim Biophys Acta* 1999; 1473: 4–8.
5. Kornfeld, R and Kornfeld, S. (1976) Comparative aspects of glycoprotein structure. *Annu Rev Biochem*, 45, 217-237
6. Drickamer, K. and Taylor, M.E.(1998) Evolving Views of protein glycosylation. *Trends Biochem Sci*, 23, 321-324
7. Rademacher, T.W., Parekh, R.B., Dwek, R.A. (1988) Glycobiology. *Annu Rev Biochem*, 57, 785-838.
8. Kornfeld, R. and Kornfeld, S.. 1985. *Annu. Rev. Biochem.* 54: 631–634.
9. Stanley P. Glycosylation mutants of animal cells. *Annu Rev Genet.* 1984; 18: 525–552
10. Freeze HH. Genetic defects in the human glycome. *Nat Rev Genet.* 2006; 7: 537–551.
11. Eklund EA, Bode L, Freeze HH. Diseases associated with carbohydrates/glycoconjugates. In: Kamerling J, Boons G-J, Lee Y. et al., editors. *Comprehensive glycoscience*. Vol. 4. Elsevier; Amsterdam: 2007. pp. 339–372.
12. Prien, Justin M., Ashline, David J., Lapadula, Anthony J., Zhang, Hailong, Reinhold, Vernon N.; *J Am Soc Mass Spectrom*, 2009
13. Dennis, J. W., Granovsky, M., et al. (1999) glycoprotein glycosylation and cancer progression. *Biochim Biophys Acta*, 1473,21-43

14. Smith, G; Leary, J.A. *J.Am. Chem.Soc.* 1996,118, 3293-3294
15. Smith G.; Perderson, S.F.; Leary, J. A. *J.org. Chem.* 1997, 62, 2152-2154.
16. Gaucher, S.P.; Leary, J.A. *Anal. Chem.*1998, 70, 3009-3254.
17. Clowers, B.H.; Dwivedi, P.; Steiner, W. E.; Hill, H.H., Jr.; Bendiak, B.; *J Am. Soc. Mass Spectrom* 2005, 16, 660-669.
18. M. D. Plasencia, D Isailovic, S.I. Merenbloom, Y. Mechref, D. E. Clemmer; *J Am Soc Mass Spectrom*, 2008, 19, 1706-1715
19. Reinhold VN. Ashline DJ, Zhang H, (2009) Unraveling the Structural Details of the Glycoproteome by ITMS. Book chapter in *Practical Aspects of Trapped Ion Mass Spectrometry* (March RE & Todd JF, Eds.). CRC Press, In press
20. Ashline, D.; Singh, S.; Hanneman, A.; Reinhold, V. *Anal. Chem.* 2005, 77, 6250-6262
21. Zhang, H.; Singh, S.; Reinhold, V. *Anal. Chem.* 2005, 77, 6263-6270
22. Lapadula, A.; Hatcher, P.; Hanneman , A.; Ashline, D.; Zhang, H.; Reihold, V. *Anal. Chem.* 2005, 77, 6271-6279
23. Prien, Justin M., Ashline, David J., Lapadula, Anthony J., Zhang, Hailong, Reinhold, Vernon N.; *J Am Soc Mass Spectrom*, 2009
24. Ciucanu, I.; Kerek, F. *Carbohydrate. Res.* 1984, 131,209-217
25. Hofmeister, F., Zeitschr, F. *Physiol. Chem.* Strassb. 14 (1889) 165.
26. Neuberger, A. *Biochem.J.* 1938, 32, 1435-1451
27. Suzuki T., Kitajima K., Emori Y., Inoue Y. and Inoue S.. *Proc. Natl. Acad. Sci. USA* 94 (1997), pp. 6244-6249.
28. Harvey D. J., Wing D.R., Kuster B., Wilson I.B.H.; *J Am Soc Mass Spectrom*, 2000,11, 564-571
29. Tai T., Yamashita K., Ogata-Arakawa M., Koide N., Murmatsu T., *J. Biol. Chem.* 250 (1975) 8569
30. Huntington James A., Stein Penelope E.; *Journal of Chromatography B*,756(2001) 189-198

31. Lattova Erika, Perreault Helene; *J Am Soc Mass Spectrom* 2004, 15, 725-735
32. Packer NH, Lawson MA, Jardine DR, Redmond JW (1998) *Glycoconj J* 15:737–747
33. Domon B.; Costello C. E.; *Glycoconjugate J* 5, 1988, 397-409
34. Douglas M. Sheeley and Vernon N. Reinhold (1998) Structural Characterization of Carbohydrate Sequence, Linkage, and Branching in a Quadrupole Ion Trap Mass Spectrometer: Neutral Oligosaccharides and N-Linked Glycans *Anal. Chem.*, 1998, 70 (14), pp 3053–3059
35. Zaia Joseph *Mass Spectrometry Review*, 2004, 23, 161-227
36. Tuch BE (2006). "Stem cells—a clinical update". *Australian Family Physician* 35 (9): 719–21.
37. Shostak S (2006). "(Re)defining stem cells". *Bioessays* 28 (3): 301–8.
38. Schöler Hans R. (2007). "The Potential of Stem Cells: An Inventory" in Nikolaus Knoepffler, Dagmar Schipanski, and Stefan Lorenz Sorgner. *Humanbiotechnology as Social Challenge*. Ashgate Publishing, Ltd. pp. 28. ISBN 0754657558.
39. Thomson J, Itskovitz-Eldor J, Shapiro S, Waknitz M, Swiergiel J, Marshall V, Jones J (1998). "Embryonic stem cell lines derived from human blastocysts". *Science* 282 (5391): 1145–7
40. Wu DC, Boyd AS, Wood KJ (2007). "Embryonic stem cell transplantation: potential applicability in cell replacement therapy and regenerative medicine". *Front Biosci* 12: 4525–35
41. Gahrton G, Björkstrand B (2000). "Progress in haematopoietic stem cell transplantation for multiple myeloma". *J Intern Med* 248 (3): 185–201.
42. Coppi P. De, Barstch G, Atala Anthony (2007). "Isolation of amniotic stem cell lines with potential for therapy". *Nature Biotechnology* 25 (5): 100-106.
43. Cairns, J. (1975) Mutation selection and the natural history of cancer. *Nature* 255, 197-200.
44. Sherley, J. L., "Asymmetric Self-Renewal: The Mark of the Adult Stem Cell," in *Stem Cell Repair and Regeneration*, eds. N. A. Habib, M. Y. Gordon, N. Levicar, L. Jiao, and G. Thomas-Black, Imperial College Press (London) pp. 21-28, 2005",

45. Rambhatla, L., Bohn, S. A., Stadler, P. B., Boyd, J. T., Coss, R. A., and Sherley, J. L. (2001) "Cellular Senescence: Ex vivo p53-Dependent Asymmetric Cell Kinetics," *J. Biomed. Biotech* 1, 27-36.
46. Rambhatla, L., Ram-Mohan, S., Cheng, J. J., Sherley, J. L. (2005) "Immortal DNA Strand Co-Segregation Requires p53/IMPDH-Dependent Asymmetric Self-Renewal Associated with Adult Stem Cells," *Cancer Research* 65, 3155-3161.
47. Merok, J. R., Lansita, J. A., Tunstead, J. R., and Sherley, J. L. (2002) "Co-segregation of Chromosomes Containing Immortal DNA Strands in Cells That Cycle With Asymmetric Stem Cell Kinetics," *Cancer Research*, 62, 6791-6795
48. Morrison, S. J. and Weissman, I. L. (1994) The long-term repopulating subset of hematopoietic stem cells in deterministic and isolatable by phenotype. *Immunity* 1, 661-673
49. 1991 Sherley, J. L. (1991) "Guanine Nucleotide Biosynthesis is Regulated by the Cellular p53 Concentration", *J. Biol. Chem.*, 266, 24815-24828.
50. Sherley, J. L., Stadler, P. B., and Johnson, D. R. (1995a) "Expression of the Wild-type p53 Antioncogene Induces Guanine Nucleotide-Dependent Stem Cell Division Kinetics", *Proc. Natl. Acad. Sci.* 92, 136-140.
51. Sherley, J. L., Stadler, P. B., and Stadler, J. S. (1995b) "A Quantitative Method for the Analysis of Mammalian Cell Proliferation in Culture in Terms of Dividing and Non-dividing Cells", *Cell Proliferation* 28, 137-144.
52. Liu, Y., Bohn, S. A., and Sherley, J. L. (1998a) "Inosine-5'-Monophosphate Dehydrogenase is a Rate-Determining Factor for p53-Dependent Growth Regulation," *Molecular Biology of the Cell* 9, 15-28.
53. Liu, Y., Riley, L. B., Bohn, S. A., Boice, J. A., Stadler, P. B., and Sherley, J. L. (1998b) "Comparison of Bax, Waf1, and IMP Dehydrogenase Regulation in Response to Wild-type p53 Expression Under Normal Growth Conditions," *J. Cellular Physiology* 177, 364-376.
54. Lanctot, Pascal M; Gage, Fred H and Varki, Ajit P. 2007 The glycans of stem cells *Current opinion in Chemical Biology* vol.11(4):373-380
55. Tero Satomaa, Annamari Heiskanen, Milla Mikkola, Cia Olsson, Maria Blomqvist, Minna Tiittanen, Olli Aitio, Anne Olonen, Jari Helin, Jukka Hiltunen, Jari Natunen, Timo Tuuri, Timo Otonkoski, Juhani Saarinen, Jarmo Laine. The N-glycome of human embryonic stem cells *BMC Cell Biology* 2009, 10:42

56. Wearne Kimberly A., Winter Harry C., O'Shea Katherine, and Goldstein Irwin J. Use of Lectins for Probing Differentiated Human Embryonic Stem Cells for Carbohydrates *Glycobiology* 2006, 16:981-990.
57. Heiskanen Annamari, Hirvonen Tia, Salo Hanna, Impola Ulla, Olonen Anne, Laitinen Anita, Tiitinen Sari, Natunen Suvi, Aitio Olli, Miller-Podraza Halina, Wuhrer Manfred, Deelder André M., Natunen Jari, Laine Jarmo, Lehenkari Petri, Saarinen Juhani. Satomaa Tero, Valmu Leena. *Glycoconj J* (2009) 26:367–384
58. Herscovics Annette, 1999, Importance of glycosidases in mammalian glycoprotein biosynthesis *Biochimica et Biophysica Acta* 1473 (1999) 96-107
59. Lowe John B., Marth Jamey D. 2003, A Genetic approach to mammalian glycan function. *Annu. Rev. Biochem.*72:643-691
60. Beyer T.A. and Hill R.L. 1982. Glycosylation pathways in the biosynthesis of nonreducing terminal sequences in oligosaccharides of glycoproteins. In *The glycoconjugates* (ed. Horowitz M. and Pigman W.), vol. III, pp. 25–45. Academic Press, New York.
61. Brisson J.-R. and Carver J.P.. 1983. The relation of three-dimensional structure to biosynthesis in the N-linked oligosaccharides *Can. J. Biochem. Cell Biol* 61: 1067-1078.
62. Brockhausen I., Hull E., Hindsgaul O., Schachter H., Shah R.N., Michnick S.W., and Carver J.P.. 1989. Control of glycoprotein synthesis: Detection and characterization of a novel branching enzyme from hen oviduct, UDP-N-acetylglucosamine: GlcNAc $\beta$ 1–6 (GlcNAc $\beta$ 1–2)Man $\beta$ -R (GlcNAc TO Man)  $\beta$ -4-N-acetylglucosaminyltransferase VI. *J. Biol. Chem.* 264: 11211-11221.
63. Harpaz N. and Schachter H.. 1980. Control of glycoprotein synthesis: Processing of asparagine-linked oligosaccharides by one or more rat liver golgi  $\alpha$ -d-mannosidases dependent on the prior action of UDP-N-acetylglucosamine: $\alpha$ -d-mannoside  $\beta$ 2-N-acetylglucosaminyltransferase I. *J. Biol. Chem.* 255: 4894-4902.
64. Lis H. and Sharon N.. 1993. Protein glycosylation: Structural and functional aspects *Eur. J. Biochem.* 218: 1-27.