Spring 5-4-2020

Characterization of Glycan and Glycoprotein by Mass Spectrometry for Therapeutic Drugs Development and Biomarker

Ding Liu

CHARACTERIZATION OF GLYCAN AND GLYCOPROTEIN BY MASS SPECTROMETRY FOR THERAPEUTIC DRUGS DEVELOPMETN AND BIOMARKER

by

DING LIU

Under the Direction of Dr. Jun Yin, PhD

ABSTRACT

Glycosylation is an essential post-translational modification of protein, which is involved in many biological processes including protein folding, immune response, protein-protein interaction and pathogens. Glycan exhibits complexity and dynamic changes in not only compositions of monosaccharides but also the linkages of them. Mass spectrometry is a powerful approach that provides a systematic and high-throughput analysis of protein glycosylation. In this dissertation, the mass spectrometry techniques have been utilized to analyze therapeutic glycoproteins and discover novel biomarkers of colon cancer. Numerous analytical chemistry techniques have been applied, including hydrophilic interaction enrichment of glycopeptide, solid phase based identification of O-glycosylation site, 18O labelling of N-glycosylation site, stepped collision energy induced glycopeptide dissociation, HCD/ETD alternative dissociation and CID

based multi stage mass spectrometry. These advanced techniques have been applied in this dissertation to comprehensively understand the glycosylation of Coagulation factor V, discover the potential glycoprotein biomarkers of colon cancer and elucidate the unrevealed structure of pneumococcal polysaccharide vaccine.

INDEX WORDS: Mass spectrometry, Glycan, Glycopeptide, Factor V, Colon cancer, Pneumococcal polysaccharide.

CHARACTERIZATION OF GLYCAN AND GLYCOPROTEIN BY MASS SPECTROMETRY FOR THERAPEUTIC DRUGS DEVELOPMETN AND BIOMARKER DISCOVERY

by

DING LIU

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

in the College of Arts and Sciences

Georgia State University

2020

CHARACTERIZATION OF GLYCAN AND GLYCOPROTEIN BY MASS SPECTROMETRY FOR THERAPEUTIC DRUGS DEVELOPMETN AND BIOMARKER DISCOVERY

by

DING LIU

Committee Chair: Jun Yin

Committee: Ning Fang

Ming Luo

Siming Wang

Lei Li

Electronic Version Approved:

Office of Graduate Services

College of Arts and Sciences

Georgia State University

May 2020

DEDICATION

I dedicate this to my family and my friends for their love, understanding, and support.

ACKNOWLEDGEMENTS

Firstly, I would like to thank my advisors Dr. Jun Yin and Dr. Peng George Wang for their mentorship during the five years of my Ph.D. journey. They lead me to enter the gate of science and establish the base of all my research. Without their guidance, I could not have gotten this achievement today.

Secondly, I would like to thank my committee Dr. Ning Fang, Dr. Ming Luo, Dr. Siming Wang and Dr. Lei Li for their time and their advices on my research and dissertation.

I would also like to thank all the members of Wang lab for their help towards my study. I owe gratitude to Dr. Cheng Ma and Dr. He Zhu for their training on Mass Spectrometry techniques, Dr. Kuan Jiang for training on glycomics techniques, Dr. Zhigang Wu and Dr. Cong Xiao for help in protein expression and purification, Dr. Jiabin Zhang for help in NMR analysis. Without the help of them, I could not become a mature analytical chemist.

Last, I would like to thank my family for their love and support!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AAs	amino acids
ACN	Acetonitrile
ADCC	antibody-dependent cellular cytotoxicity
AGT	Angiotensinogen
APC	activated protein C
BCA	icinchoninic acid
BSA	Bovine serum albumin
CA	carbohydrate antigen
CEA	carcinoembryonic antigen
CH2	constant heavy 2
CO3	Complement C3
CO7	complement component C7
CPS	capsular polysaccharides
CRC	Colorectal cancer
DTT	Dithiothreitol
ECM	extracellular matrix
ELSD	evaporative light scattering detector
ESL	experimental spectral library
ETD	electron-transfer dissociation
EXoO	site-specific extraction of O-linked glycopeptide
FA	formic acid
FASP	Filter-aided Sample Preparation

FDR	false discovery rate
FFP	fresh frozen plasma
FTMS	Fourier Transform Mass Spectrometry
FV	Coagulation Factor V
FVIII	factor VIII
FXIII	Coagulation factor XIII
GAGs	glycosaminoglycans
Gal-3BP	Galectin-3-binding protein
Glu-C	endoproteinase Glu-C
GO	Gene Ontology
HCC	hepatocellular carcinoma
HCD	higher collision energy dissociation
HILIC	hydrophilic interaction chromatography
HSQC	heteronuclear single quantum coherence
IAA	iodoacetamide
IgA	Immunoglobulin A
IgG1	Immunoglobulin G1
IgM	Immunoglobulin M
LCA	Lens culinaris agglutinin
LMWHs	low molecular weight heparins
LPS	lipopolysaccharide
m/z	mass-to-charge
MRM	multiple reaction monitor

NCE	normalized collision energy
Neu5Gc	N-glycolylneuraminic acid
NH4HCO3	ammonium hydrogen carbonate
OFAT	one factor at a time
PAI-1	plasminogen activator inhibitors 1
PLTP	Phospholipid transfer protein
Pn8P	pneumococcal type 8 polysaccharide
PNGase F	Peptide-N-glycosidase F
PTM	posttranslational modification
RID	refractive index detector
RSD%	Relative standard deviation error
RT	retention time
SEC	size exclusion chromatography
Sia6LacNAc	N-acetyl lactosamine
sLex	Sialyl Lewis x
SNCE	stepped normalized collision energy
ST6Gal-I	α2,6-sialyltransferase I
TFA	Trifluoroacetic acid
uPA	urokinase plasminogen activator
WR	working reagent
ZIC-HILIC	zwitterionic hydrophilic interaction chromatography

1 CHAPTER 1. INTRODUCTION TO GLYCOMICS AND GLYCOPROTEOMICS

For many years, researchers have been focusing on exploration of the biological information of DNAs, RNAs and proteins. In fact, small molecule metabolites such as glycan also play important role in biological system. Studying the structure and distribution of glycan and glycoconjugate is essential to understand the biological function of these molecules. In this chapter, the method and application of glycomics and glycoproteomics are briefly introduced.

1.1 Introduction to glycan and glycosylation

Glycans or carbohydrates, including monosaccharide, oligosaccharide and polysaccharides, are a series of molecules that widely distributed in nature. There are nine common monosaccharides in the organism of vertebrates, these common monosaccharides construct most of the oligosaccharides.[1] (Figure 1.1A) Starting from initial monosaccharide, glycan can be extended through glycosidic linkages to oligosaccharide and polysaccharide. As a posttranslational modification (PTM), glycosylation is an enzymatic reaction that oligosaccharide covalently binds to proteins through N- or O- linkages. N-linked oligosaccharide (N-glycan) is an oligosaccharide attached to asparagine of peptide backbone with its innermost GlcNAc. The corresponding peptide sequence is Asn-X-Ser/Thr, and the glycan contains core pentasaccharide. N-glycan can be divided into three main types, high-mannose type, hybrid type and complex type.[1] (Figure 1.1B) O-linked oligosaccharide (O-glycan) is an oligosaccharide attached to serine or threonine of polypeptide through GalNAc. O-glycan usually contains four core structures, which can be further extended to more complex structure.[1] (Figure 1.1C) N-glycan plays important roles in protein folding, signaling, transporting and cell-cell interaction.[2] O-glycan can influence protein confirmation and stability, it can also act as recognition epitopes in immune system.[3]

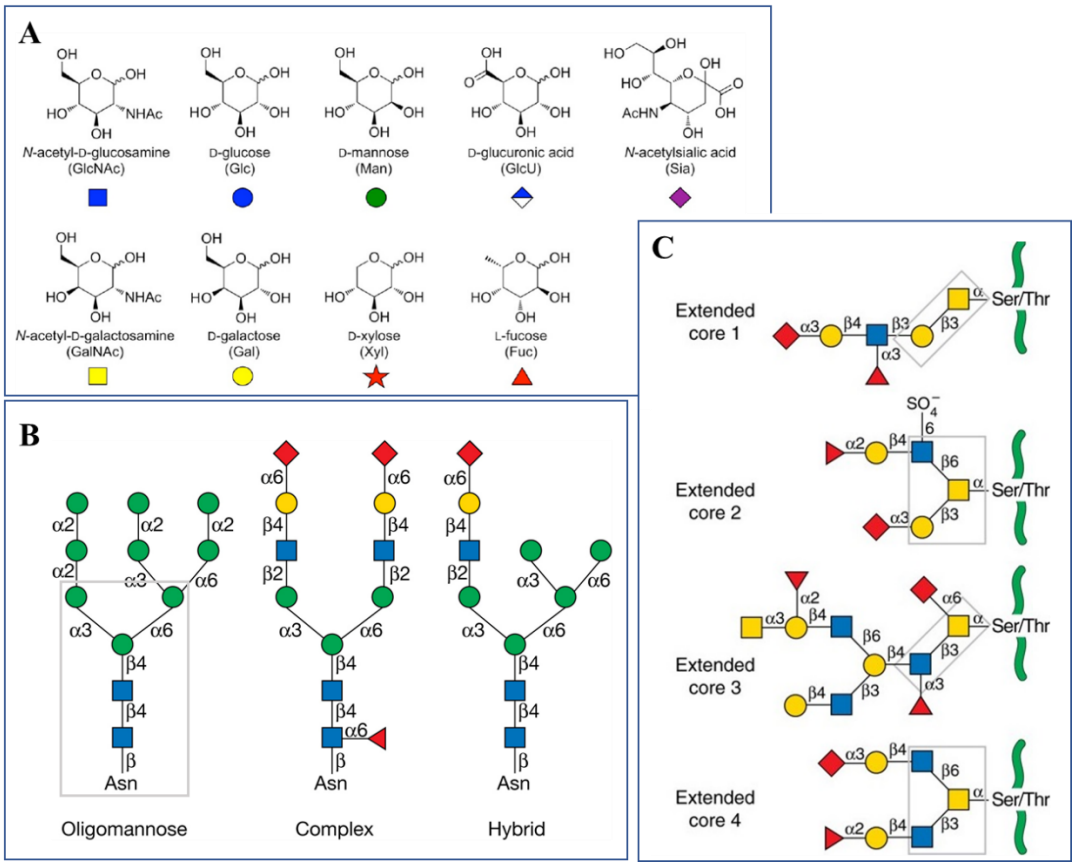


Figure 1.1 A. Nine common mononsacharides in vertebrates; B. N-glycan core structure and main types; C. O-glycan core structures and main types.

Differently from oligosaccharide, polysaccharide is long chain polymeric carbohydrates. There are several types of polysaccharides, including plan polysaccharide, animal glycosaminoglycans (GAGs) and bacterial polysaccharide. Bacterial polysaccharides, including lipopolysaccharides (LPSs), capsular polysaccharides (CPSs), and exopolysaccharides (EPSs), are antigens that can cause strong immune response. Polysaccharide has been widely applied as drugs or drugs delivery. For example, hyaluronic acid has been used for joint disorders and eye surgeries[4], pneumococcal polysaccharide has been developed as vaccine and vaccine delivery.[5] Since the increasing recognization of glycan's importance, glycomics and glycoproteomics have been developed to study glycan and glycoconjugate in recent decades.

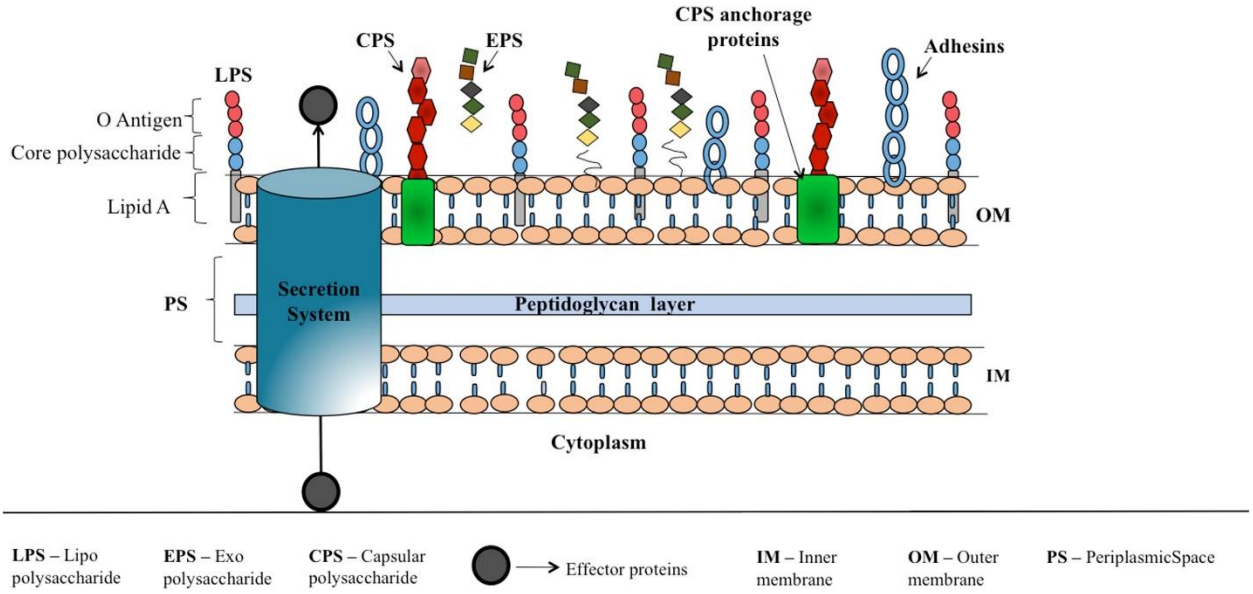


Figure 1.2 Bacterial polysaccharides distribute on surface of bacteria, including LPS, CPS and EPS.

1.2 Glycan and glycosylation analysis by Mass Spectrometry

Glycomics is designed in recent decades to identify and characterize the glycan from cells or tissues.[6] Distinguished from gene and protein, glycan is not template driving synthesized, which increase the diversity and complexity of the structures. Until the development of Mass Spectrometry technique in this century, glycomics and glycoproteomics start to attract researchers' attention. Glycan can be studied solely after detached from proteins, which is glycomics. N-glycan can be released from glycoprotein by PNGase F enzymatic digestion, O-glycan can be released by beta-elimination chemically.[7, 8] Both of N- and O- glycan generally need to be derivatized to improve the signal in mass spectrometer.[9] Polysaccharide is hard to be ionized because of its large size, while enzymatic digestion of polysaccharide to smaller fragments is always applied to solve this problem.[10]

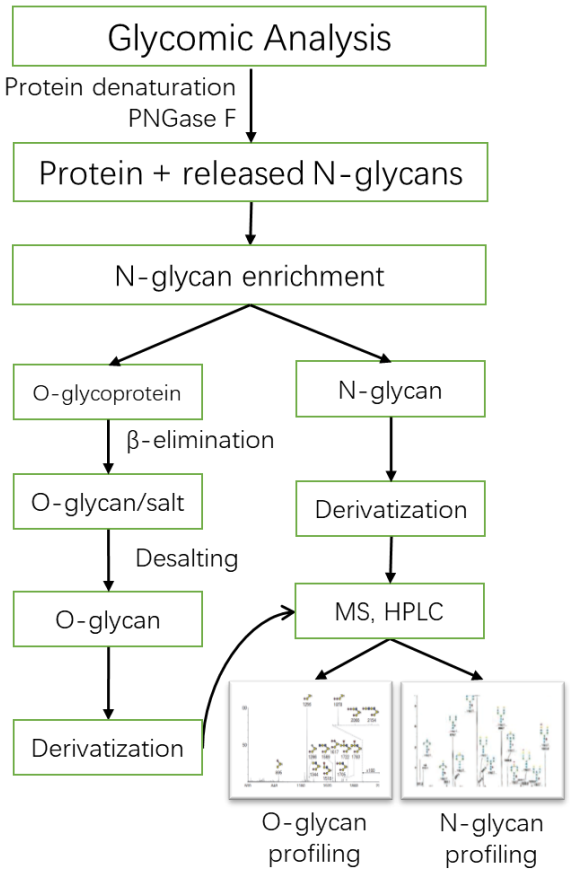


Figure 1.3 General process of glycomics to study N- and O- glycan released from glycoprotein.

Glycan can also be studied when covalently linked with protein as PTM, which is glycoproteomics. Glycopeptide can be enriched through several methods, including hydrazide chemistry enrichment, lectin affinity chromatography, hydrophilic interaction chromatography (HILIC).[11-13] Then glycosylation site analysis and intact-glycopeptide analysis can be performed based on the enriched glycopeptide. N-glycosylation sites are identified by labeled N-glycosylated asparagine by ^{18}O. The asparagine would be transferred to aspartic acid when digestion with PNGase F, the oxygen on the terminal of carboxyl group would be labeled when digested condition is $H_2{}^{18}O$.[14] By using this digestion method, N-glycosylation sites would have +2.998Da difference with asparagine, which can be easily identified by mass spectrometry. O-glycosylation site

analysis method has been recently developed.[15] Peptide can be covalently bind to aldehyde-activated beads, then only O-glycopeptide can be site specifically cleaved from C terminal of Ser/Thr with O-glycan on it. O-glycopeptide can thereby enriched and analyzed by mass spectrometry.[15]

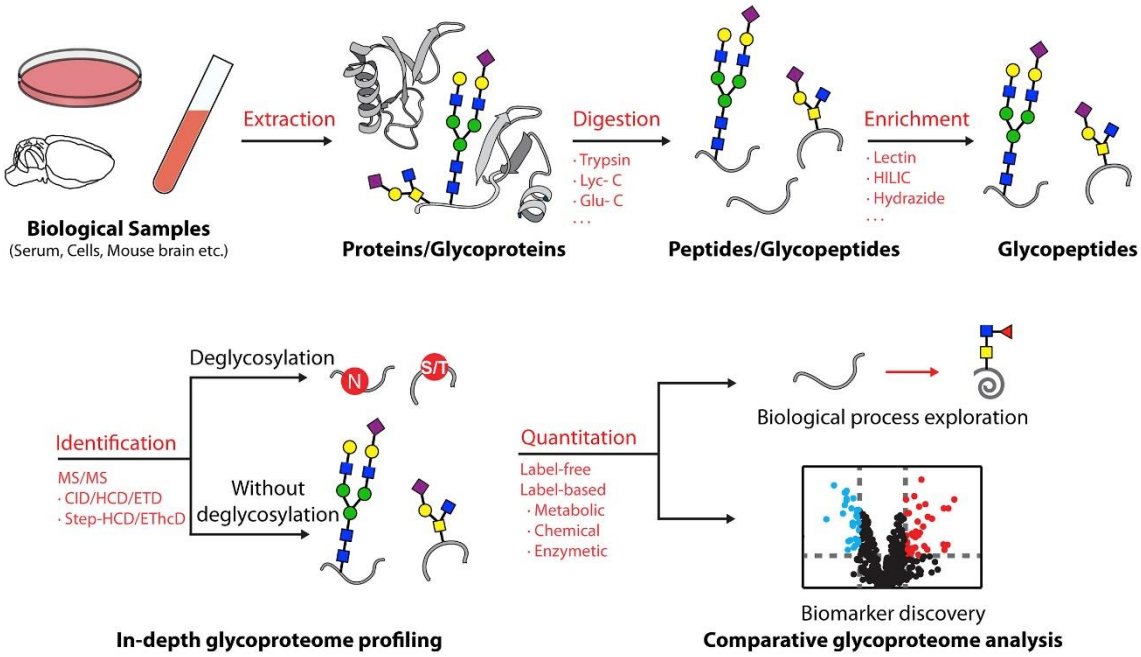


Figure 1.4 General workflow for glycoproteomics to study glycopeptide from biological samples.

Intact-glycopeptide analysis is achieved by recent developed tandem MS dissociation method and database search engine. Intact-glycopeptide was lack of appropriate dissociation method until recent. Since different higher collision energy dissociation (HCD) collision energies could produce more fragments of glycopeptide, step-HCD has been proved to significantly improve the overall identification of intact-glycopeptide.[16] Complementary to HCD, electron-transfer dissociation (ETD) enables the cleavage of the peptide backbone by maintaining the glycan-peptide linkage, and the fragmentation information can help to analyze peptide sequences. Thus, EThcD can greatly facilitate the identification of both glycoforms and glycosites.[17] Besides

the dissociation method, new search engines have been developed for intact-glycopeptide, including GPQuest,[18] pGlyco 2.0[19] and Byonic.[20] These new software make large-scale intact-glycopeptide analysis achievable.

Since glycan can be affected by environmental factors, glycomics and glycoproteomics have been applied to identify disease biomarkers and monitor drug efficiency.[1] For example, Shah et al combined proteomics and glycoproteomics to study the mechanism of castration resistance for androgen-deprivation therapy in prostate cancer.[21] Simeone and coworkers utilized *Lens culinaris* agglutinin (LCA) to enrich core-fucosylated glycoproteins, and found 8 core-fucosylated glycopeptide significantly different in pancreatic cancer and controls.[22] The glycosylation of therapeutic protein—FVIII from different sources was completely studied and compared to help solving the inhibition problem in clinical therapy cases.[23]

1.3 Scope

Numerous therapeutic proteins and disease biomarkers are glycoproteins, study of glycan is necessary to understand the functional mechanism of these glycoproteins. Mass spectrometry has been a robust tool to study glycan and glycoprotein. Therefore, the aim of this dissertation is the application of glycomics and glycoproteomics to therapeutic drugs and biomarkers discovery.

In **chapter 2**, the glycosylation patterns of human plasma derived FV was comprehensively mapped through combining multi-enzyme digestion, hydrophilic interaction chromatography enrichment of glycopeptides, and HCD/ETD alternated fragmentation mass spectrometry analysis. A data analysis strategy for intact glycopeptide sequencing is designed and applied in this work.

In **chapter 3**, an integrated glycoproteomics study of colon cancer patient's serum was performed. Glycosylation sites were labelled by ^{18}O, and an experimental glycopeptide spectra

library was built. The library was further applied to identify intact-glycopeptide under stepped normalized collision energy. Overall glycan pattern of glycoprotein in colon cancer serum was observed. Several glycoproteins and N-glycans were selected as potential biomarker of colon cancer.

In **chapter 4**, heparinase was firstly found to depolymerize type 8 pneumococcal polysaccharide vaccine. Several polysaccharide lyase and hydrolase were reacted with the polysaccharide, and heparinase was surprisingly able to digest it. The reaction condition was optimized, and the digested product was purified to oligosaccharide. Then the oligosaccharide was characterized by multiple stage mass spectrometry and two-dimensional NMR.

1.4 Reference

1. Wiederschain, G. Y., Essentials of glycobiology. Springer Science & Business Media: 2009.
2. Helenius, A.; Aebi, M. J. S., Intracellular functions of N-linked glycans. **2001,** *291* (5512), 2364-2369.
3. Patsos, G.; Corfield, A., *O-Glycosylation: structural diversity and function*. Wiley-VCH: Weinheim, Germany: 2009.
4. Modawal, A.; Ferrer, M.; Choi, H. K.; Castle, J. A., Hyaluronic acid injections relieve knee pain. In *Database of Abstracts of Reviews of Effects (DARE): Quality-assessed Reviews [Internet]*, Centre for Reviews and Dissemination (UK): 2005.
5. Kenneth, J. R.; Ray, C. G. J. V., Sherris medical microbiology. **2004**.
6. Hart, G. W.; Copeland, R. J. J. C., Glycomics hits the big time. **2010,** *143* (5), 672-676.
7. Fukuda, M. J. C. p. i. m. b., β‐Elimination for Release of O‐GalNAc‐Linked Oligosaccharides from Glycoproteins and Glycopeptides. **1995,** *31* (1), 17.15. 4-17.15. 9.
8. Vanhooren, V.; Laroy, W.; Libert, C.; Chen, C. J. B., N-Glycan profiling in the study of human aging. **2008,** *9* (5), 351.
9. Ruhaak, L.; Zauner, G.; Huhn, C.; Bruggink, C.; Deelder, A.; Wuhrer, M. J. A.; chemistry, b., Glycan labeling strategies and their use in identification and quantification. **2010,** *397* (8), 3457-3481.
10. Wang, Z.; Zhang, T.; Xie, S.; Liu, X.; Li, H.; Linhardt, R. J.; Chi, L. J. C. p., Sequencing the oligosaccharide pool in the low molecular weight heparin dalteparin with offline HPLC and ESI–MS/MS. **2018,** *183*, 81-90.
11. Huang, G.; Sun, Z.; Qin, H.; Zhao, L.; Xiong, Z.; Peng, X.; Ou, J.; Zou, H. J. A., Preparation of hydrazine functionalized polymer brushes hybrid magnetic nanoparticles for highly specific enrichment of glycopeptides. **2014,** *139* (9), 2199-2206.

12. Badr, H. A.; AlSadek, D. M.; Darwish, A. A.; ElSayed, A. I.; Bekmanov, B. O.; Khussainova, E. M.; Zhang, X.; Cho, W. C.; Djansugurova, L. B.; Li, C.-Z. J. E. r. o. p., Lectin approaches for glycoproteomics in FDA-approved cancer biomarkers. **2014,** *11* (2), 227-236.
13. Hägglund, P.; Bunkenborg, J.; Elortza, F.; Jensen, O. N.; Roepstorff, P. J. J. o. p. r., A new strategy for identification of N-glycosylated proteins and unambiguous assignment of their glycosylation sites using HILIC enrichment and partial deglycosylation. **2004,** *3* (3), 556-566.
14. Küster, B.; Mann, M. J. A. c., 18O-labeling of N-glycosylation sites to improve the identification of gel-separated glycoproteins using peptide mass mapping and database searching. **1999,** *71* (7), 1431-1440.
15. Yang, W.; Ao, M.; Hu, Y.; Li, Q. K.; Zhang, H. J. M. s. b., Mapping the O‐glycoproteome using site‐specific extraction of O‐linked glycopeptides (EXoO). **2018,** *14* (11).
16. Cao, Q.; Zhao, X.; Zhao, Q.; Lv, X.; Ma, C.; Li, X.; Zhao, Y.; Peng, B.; Ying, W.; Qian, X. J. A. c., Strategy integrating stepped fragmentation and glycan diagnostic ion-based spectrum refinement for the identification of core fucosylated glycoproteome using mass spectrometry. **2014,** *86* (14), 6804-6811.
17. Glover, M. S.; Yu, Q.; Chen, Z.; Shi, X.; Kent, K. C.; Li, L. J. I. J. o. M. S., Characterization of intact sialylated glycopeptides and phosphorylated glycopeptides from IMAC enriched samples by EThcD fragmentation: toward combining phosphoproteomics and glycoproteomics. **2018,** *427*, 35-42.
18. Toghi Eshghi, S.; Shah, P.; Yang, W.; Li, X.; Zhang, H. J. A. c., GPQuest: a spectral library matching algorithm for site-specific assignment of tandem mass spectra to intact N-glycopeptides. **2015,** *87* (10), 5181-5188.
19. Liu, M.-Q.; Zeng, W.-F.; Fang, P.; Cao, W.-Q.; Liu, C.; Yan, G.-Q.; Zhang, Y.; Peng, C.; Wu, J.-Q.; Zhang, X.-J. J. N. c., pGlyco 2.0 enables precision N-glycoproteomics with comprehensive quality control and one-step mass spectrometry for intact glycopeptide identification. **2017,** *8* (1), 1-14.
20. Bern, M.; Kil, Y. J.; Becker, C. J. C. p. i. b., Byonic: advanced peptide and protein identification software. **2012,** *40* (1), 13.20. 1-13.20. 14.
21. Shah, P.; Wang, X.; Yang, W.; Eshghi, S. T.; Sun, S.; Hoti, N.; Chen, L.; Yang, S.; Pasay, J.; Rubin, A. J. M.; Proteomics, C., Integrated proteomic and glycoproteomic analyses of prostate cancer cells reveal glycoprotein alteration in protein abundance and glycosylation. **2015,** *14* (10), 2753-2763.
22. Tan, Z.; Yin, H.; Nie, S.; Lin, Z.; Zhu, J.; Ruffin, M. T.; Anderson, M. A.; Simeone, D. M.; Lubman, D. M. J. J. o. p. r., Large-scale identification of core-fucosylated glycopeptide sites in pancreatic cancer serum using mass spectrometry. **2015,** *14* (4), 1968-1978.
23. Canis, K.; Anzengruber, J.; Garenaux, E.; Feichtinger, M.; Benamara, K.; Scheiflinger, F.; Savoy, L. A.; Reipert, B.; Malisauskas, M. J. J. o. T.; Haemostasis, In‐depth comparison of N‐glycosylation of human plasma‐derived factor VIII and different recombinant products: from structure to clinical implications. **2018,** *16* (8), 1592-1603.

2 CHAPTER 2. COMPREHENSIVE N- AND O-GLYCOSYLATION MAPPING OF HUMAN COAGULATION FACTOR V

2.1 Abstract

Coagulation Factor V (FV), a multi-domain glycoprotein, is an essential cofactor in the blood clotting cascade. The deficiency or dysfunction of FV causes Owren's disease, a rare bleeding disorder that results in poor clotting after an injury or surgery. The only treatment for the disease is infusions of fresh frozen plasma and blood platelets. Glycosylation affects the biological activity, pharmacokinetics, immunogenicity, and in vivo clearance rate of proteins in the plasma. The glycan profile of FV, as well as how it affects the activity, stability, and immunogenicity remain unknown. In this study, we comprehensively mapped the glycosylation patterns of human plasma derived FV through combining multi-enzyme digestion, hydrophilic interaction chromatography enrichment of glycopeptides, and alternated fragmentation mass spectrometry analysis. A total of 57 unique N-glycopeptides and 51 O-glycopeptides were identified, which were categorized into 40 N-glycan and 17 O-glycan compositions. Such glycosylation details are fundamental for future functional studies and therapeutics development. In addition, the established methodology can be readily applied to analyze glycosylation patterns of proteins with over 2000 amino acids.

2.2 Introduction

Human coagulation factor V (FV), also known as proaccelerin, is an essential glycoprotein involved in the clotting cascade. FV is expressed primarily in the liver, the 2224 amino acids (AAs) protein (including a 28 AAs signaling peptide) consists of six domains in the following primary sequence, A1, A2, B, A3, C1, and C2, similar to that of factor VIII (FVIII). FV and FVIII share approximately 40% AA sequence homology within A and C domains.[1, 2] Upon proteolytic activation by either thrombin or FVa, the B domain, which is unnecessary for procoagulant activity,

is cleaved from FV. FV is involved in converting prothrombin into thrombin, a vital step in the blood clotting process. Clinically, this rare bleeding disorder caused by FV deficiency is manifested as a prolonged blood clotting problem after an injury or surgery. Currently, fresh frozen plasma (FFP) is the standard therapeutic option.[3] However, certain FV-deficient patients developed inhibitors after receiving FFP.[4, 5] Plasma-derived FV concentrate were tested as an alternative but not widely used given difficulties in isolation.[6] Additionally, FV deficiency may occur at the same time as FVIII deficiency, which is considered as a separate blood clotting disorder and manifests a more severe hemophiliac phenotype.[7]

Glycosylation has been reported to play critical roles in protein folding, stability, macromolecular interactions, and activity.[8-12] Although the AA sequence within the B domain of FV apparently diverges between human and bovine genes, large numbers of glycosylation sites are conserved, suggesting that glycosylation may provide a regulatory role for both the expression and activation of FV.[13] On the one hand, glycosylation affects the function of FV and is important for interaction between FV and other factors involved in the clotting cascade. For example, two isoforms of plasma FV that differ in glycosylation at Asn2209 in the C2 domain, FV1 (~33%) is N-glycosylated at this site whereas FV2 (~67%) is not, showed different functionality.[14] The activated FVa2 has a higher affinity for phospholipids than FVa1, thus has higher procoagulant potency. This had been shown to interfere with the formation of FVa-FXa complex.[15] In addition, total removal of N-glycans and terminal sialic acid on FV by treatment with glycosidases resulted in increased activated protein C (APC) sensitivity, suggesting that glycosylation plays a role in APC-catalyzed cleavage and inactivation of FV.[16] On the other hand, a FV missense mutation Ile359Thr, which creates an additional N-glycosyltion at Asn 357, could significantly reduce the cleave at Arg306, thus affect anticoagulation.[17] Similarly, FV mutants that contain unnatural N-

glycosyaltion on position 495, 539, 680, or 1710 displyed attenuated FXa binding, which could be restored by inhibition of N-glycan expression, again implies the roles carbohdrates might play in mutants FV-FX binding.[18] Unlike Asn2209, these positions locate on A2 and A3 domians of FV, it is unknown whether these domains contain glycosylation and do they play functional roles. Nevertheless, a comprehensive knowledge of the glycosylation of FV could provide fundamentals for understanding the roles and mechanism of FV in clotting, as recent reported on glycan profiling of FVIII.[19-22] Additionally, such information and a method to profile FV glycosyaltion could of great importance in the development of FV-based pharmaceutics, as altered glycosylation was reported to affect the immunogenicity of glycoprotein durgs, e.g., non-human glycan structures (e.g., alpha-gal, or N-glycolylneuraminic acid) on recombinant FVIII (rFVIII) proteins werer reported to increase their intrinsic immunogenicity.[23, 24] To date, a systemic analysis of FV glycosylation is still missing.

In this work, we developed and applied an integrated approach for systematic glycan analysis of plasma-derived FV (pdFV). Glu-C and Trypsin were used for peptide digestion, and glycopeptide enrichment was performed by zwitterionic hydrophilic interaction chromatography (ZIC-HILIC). Alternated higher-energy collisional dissociation (HCD) and electron-transfer dissociation (ETD) fragmentation were employed to integrate N- and O-glycopeptide sequencing in one MS analysis.[25] [18]O-labeling of N-glycosylation sites and site-specific extraction of O-linked glycopeptide (EXoO) was also applied to localize exact N- and O-glycosylation sites.[26, 27] As results, 108 unique glycopeptides comprising 12 N-glycosylation sites and 26 O-glycosylation sites were identified with the simultaneous determination of peptide sequences and glycoform compositions.

2.3 Experimental

2.3.1 Chemicals and materials

Human pdFV was purchased from Haematologic Technologies (Essex Junction, VT). Trypsin (Sequencing grade modified) and endoproteinase Glu-C (Glu-C) were supplied by Promega (Madison, WI). PNGase F was obtained from New England Biolabs (Beverly, MA). Sodium cyanoborohydride ($NaCNBH_3$), Formic acid (FA), ammonium hydrogen carbonate (NH_4HCO_3), and HPLC grade acetonitrile (ACN) were purchased from Sigma-Aldrich (St. Louis, MO). Dithiothreitol (DTT) and iodoacetamide (IAA) were obtained from Thermo Fisher (Waltham, MA). The 10-kDa Microcon centrifugal filter devices were purchased from Millipore (Bedford, MA). ZIC-HILIC material was bought from SeQuant (Umea, Sweden); the 3M Empore C8 disk was purchased from 3M Bioanalytical Technologies (Minnesota, USA); 10 μL extended pipet tips were from Axygen (California, USA). The OpeRATOR/SialEXO kit was from Genovis, Inc (Cambridge, MA). All chemicals used in the preparation of buffers and solutions were of analytical reagent grade or better.

2.3.2 Sample preparation and enzymatic digestion

Excipients in the pdFV sample were removed by passing through a 10-kDa Microcon centrifugal filter device. The buffer was changed to 50 mM NH_4HCO_3 (pH 7.8), and the protein was concentrated to about 1 μg/μL. Filter-aided Sample Preparation (FASP) was used to prepare MS analysis samples.[28] Briefly, the above solution was incubated with 50 mM DTT at 95°C for 10 min, and then centrifuged at 16,000 × g for 10 min. The resulting supernatant was mixed with 200 μL of 8 M urea in 100 mM Tris-HCl buffer, pH 8.5 (UA solution). The mixed sample was loaded into a 30 kDa Microcon filtration device and centrifuged at 14,000 × g until the remaining volume was less than 20 μL. The concentrate was then diluted in the filter device with 200 μL UA

solution and centrifuged twice. Subsequently, the concentrate was mixed with 100 μL of 50 mM IAA in the UA solution, incubated in darkness at room temperature for 30 min, and then centrifuged for 20 min. The concentrate was subsequently diluted with 200 μL of UA solution and concentrated again. The last step was repeated twice. The sample was then diluted with 100 μL 40 mM NH_4HCO_3 and concentrated twice. After concentrating, 8 μg trypsin or 8 μg Glu-C in 100 μL 40 mM NH_4HCO_3 (pH 7.8) or 100 μL 50 mM NaH_2PO_4 (pH 7.5) was added for overnight digestion at 37°C. The resulting peptides were collected by centrifuging the filter units with 50 μL NH_4HCO_3 (pH 7.8) for 20 min. This step was repeated 3 times. The final concentration of peptides was determined by UV-spectrometry (Nanodrop, Thermo) using an extinction coefficient of 1.1 for 0.1% (g/L) solution at 280 nm.

2.3.3 Glycopeptides enrichment

Homemade HILIC SPE micro-tips were used for intact glycopeptide enrichment. A small piece of C8 membrane was taken from a 3M Empore C8 disk and pushed into the end of a 10 μL extended pipet tip using a blunt needle, 2 mg ZIC-HILIC material was then packed as previously described.[29] The HILIC SPE micro-tip was washed with 100 μL of binding buffer (80 % ACN/5 % FA) twice. For sample loading, 100 μL of binding buffer was added to the dried sample and loaded onto the micro-tip three times to allow the binding of glycopeptides. The micro-tip was then washed with the binding buffer to remove non-glycopeptides, and glycopeptides were eluted with 100 μL of elution buffer (0.5% FA) twice. The flow-through was vacuum-dried and stored at -20 °C until use. Glycopeptides (100 μg) enriched from proteolytic digested peptides were directly injected to MS for intact glycopeptide analysis.

2.3.4 N-glycosite analysis

HILIC enriched glycopeptides (100 μg) were dissolved with 50 μL of 50 mM NH_4HCO_3 in $H_2{}^{18}O$. PNGase F was then added and the solution was incubated at 37 °C for 12 h to remove N-glycans. Deglycosylated peptides were desalted using Ziptip C18 (ZTC18S096, Merck Millipore Ltd), vacuum-dried and preserved at -20 °C for mass spectrometry analysis.

2.3.5 O-glycosite analysis

Approximately 200 μg of digested peptides were dissolved in 50 mM NaH_2PO_4 (pH 7.5). The EXoO approach was used to identify O-glycosylation sites as previously reported.[27] Briefly, 200 μL Aminolink resin (Pierce, Rockford, IL) was incubated with the peptides in 50 mM $NaCNBH_3$ and 50 mM NaH_2PO_4 (pH 7.5) at room temperature overnight. The resin was then washed in a spin column and blocked by 50 mM $NaCNBH_3$ and 1M Tris-HCl (pH 7.4) at room temperature for 30 mins. After washing 3 times, the O-glycopeptides were released by OpeRATOR and SialEXO (1 unit/1 μg peptides each enzyme) in 20 mM Tris-HCl (pH 6.8) at 37°C for 15 h. The released O-glycopeptides were collected, desalted by a C18 SPE cartridge (Thermo Scientific), and dried for MS analysis.

2.3.6 LC−MS/MS analysis of glycopeptides

Experiments were performed on an LTQ-Orbitrap Elite mass spectrometer (MS) equipped with EASY-spray source and nano-LC UltiMate 3000 high-performance liquid chromatography system (Thermo Fisher). An EASY-Spray PepMap C18 Column (length, 15 cm; particle size, 3 μm; pore size, 100 Å; Thermo Fisher) was used for separation. Separation was achieved with a linear gradient from 3% to 40% solvent B for 30 min at a flow rate of 300 nL/min (mobile phase A, 2% ACN, 98% H_2O, 0.1% FA; mobile phase B, 80% ACN, 20% H_2O, 0.1% FA). The LTQ-Orbitrap Elite was operated in data-dependent mode, and ten most intense ions in MS1 were

subjected to CID in the ion trap analyzer for deglycosylated peptide analysis, or alternated HCD and ETD fragmentation for intact glycopeptide analysis. The Orbitrap MS acquired a full-scan survey (m/z range from 375 to 1,500; automatic gain control target, 106 ions; resolution at m/z 400, 60,000; maximum ion accumulation time, 50 ms). For CID-MS, the default charge state was 3, the isolation width was m/z 3.0, normalized collision energy (NCE) was 35%, activation Q was 0.25, and activation time was 5.0 ms. For alternated HCD and ETD-MS, the Orbitrap analyzer acquired HCD fragment ion spectra with a resolution of 15,000 at m/z 400 (automatic gain control target, 10,000 ions; maximum ion accumulation time, 200 ms). The LTQ analyzer acquired ETD fragment ion (automatic gain control target, 5,000 ions; maximum ion accumulation time, 100 ms). In this acquired method, the five most intense ions were fragmented by HCD and ETD. The MS/MS scan model was set as the centroid. Other conditions used were S-lens RF level of ~60%, ion selection threshold of 50,000 counts for HCD.

2.3.7 Data analysis

Data collected by the N-glycosite mapping experiment was processed with Proteome Discoverer 1.4 (Thermo Fisher Scientific). Peptide fragments were matched against the FV protein sequence (UniProtKB entry P12259 or FA5_HUMAN), where Iodoacetamide on Cys was used as static modification, oxidation of Met and ^{18}O labeling of Asn (m = 2.9848) were used as a dynamic modification. The mass tolerance was set at 10 ppm for precursor ions and 0.5 Da for-product ions. Trypsin and Glu-C were chosen for the enzyme, and two missed cleavages were allowed. A false discovery rate (FDR) of 1% was estimated and applied at the peptide level. For intact N- and O-glycopeptide analysis, all spectra generated from glycopeptides were selected and combined into an mgf file with an in-house software based on the presence of HexNAc fragment ions of m/z 126.0549, 138.0549, 144.0655, 168.0760 and 204.0866 from HCD spectra. Theoretical masses of

all possible peptides were calculated by a computer script and manually validated by PeptideMass (http://web.expasy.org/peptide_mass).[30] The online software GlycoMod (http://web.expasy.org/glycomod/) was utilized for intact N- and O-glycopeptide data interpretation. Alternated HCD and ETD fragmentation strategies were selected to analyze intact glycopeptide. To obtain precise glycopeptide sequences, we developed a data analysis strategy for intact glycopeptide sequencing. Firstly, spectra of potential N- and O-glycopeptides were selected according to diagnostic ions from a series of HexNAc fragments. Secondly, all selected spectra were analyzed with GlycoMod, and further verified manually according to Y1 ions ([peptide+GlcNAc]n+, n=1,2,3…n) and Y0 ions ([peptide]n+, n=1, 2, 3…n) from HCD-MS/MS spectra. Finally, HCD-associated ETD data were used for backbone sequencing. The workflow of MS data analysis is shown in Figure 2.1.

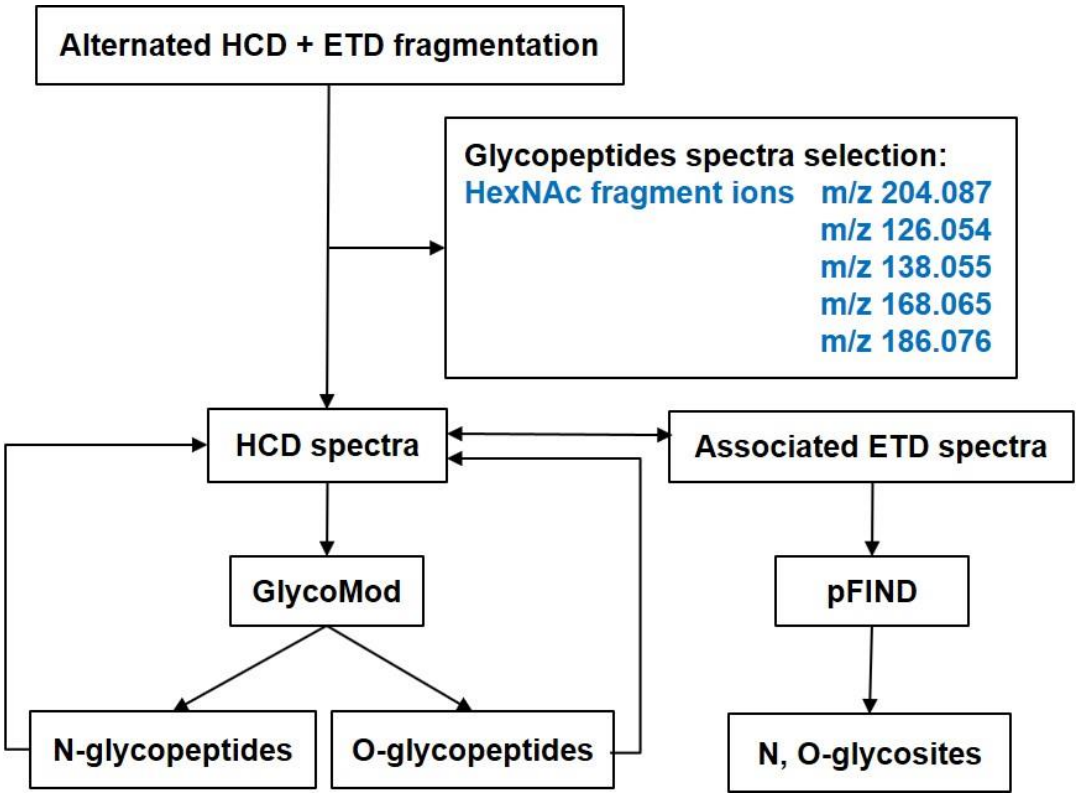


Figure 2.1 The workflow of HCD- and ETD-MS/MS analysis.

2.4 Results and Discussion

2.4.1 Glycopeptide fragmentation strategy

For comprehensive and simultaneous analysis of *N*- and *O*-glycopeptides, an alternated HCD and ETD fragmentation strategy were applied by using an Orbitrap Elite MS/MS. The parameters of alternated HCD and ETD fragmentation for CF-glycopeptides fragmentation were optimized in our previous study.[31] The HCD technique surmounts the low mass cutoff problem of ion trap fragmentation, and dramatically improves the quality of MS/MS spectra given the high accuracy at both precursor mass and fragment levels.[32] From HCD spectra of intact glycopeptides, the fragment ion HexNAc attached to the peptide rupture into a series of *m/z* 126.0549, 138.0549, 144.0655, 168.0760, and 204.0866. Meanwhile, some oligosaccharide fragments are also observable, such as Hex-HexNAc (*m/z* 366.1395) and Hex-Hex-HexNAc (*m/z* 528.1923). All diagnostic ions were used for selecting and picking out the spectra of glycopeptides. Another noteworthy factor is that *N*- and *O*-glycopeptide types can be distinguished from the fragmentation pattern observed in HCD spectra. In fact, *N*-glycans are linked through GlcNAc to Asn residues with an amide linkage, whereas mucin-type *O*-linked carbohydrate attachments to proteins involve a covalent linkage between the GalNAc and amino acid Ser or Thr with an *O*-glycosidic bond. In the fragmentation process, glycosidic bonds are easier to break than peptide bonds.[33] It means that *N*-glycopeptides often dissociate between the two innermost GlcNAc residues, generating a typical Y1 ion in MS/MS spectra,[34] while *O*-glycopeptides typically lose the entire glycan and generate intact peptide ion (Y0) in MS/MS spectra (Figure 2.2). The diagnostic peaks of Y1 and Y0 were used to verify the sequence of glycopeptides by comparing them with theoretical masses of all possible peptides. Aside from HCD, another widely utilized fragmentation for glycopeptide analysis is ETD.[35, 36] Complementary to HCD, ETD enables the cleavage of the peptide backbone

by maintaining the glycan-peptide linkage, and the fragmentation information can help to analyze peptide sequences. Thus, combined fragmentation of HCD and ETD can greatly facilitate the identification of both glycoforms and glycosites.

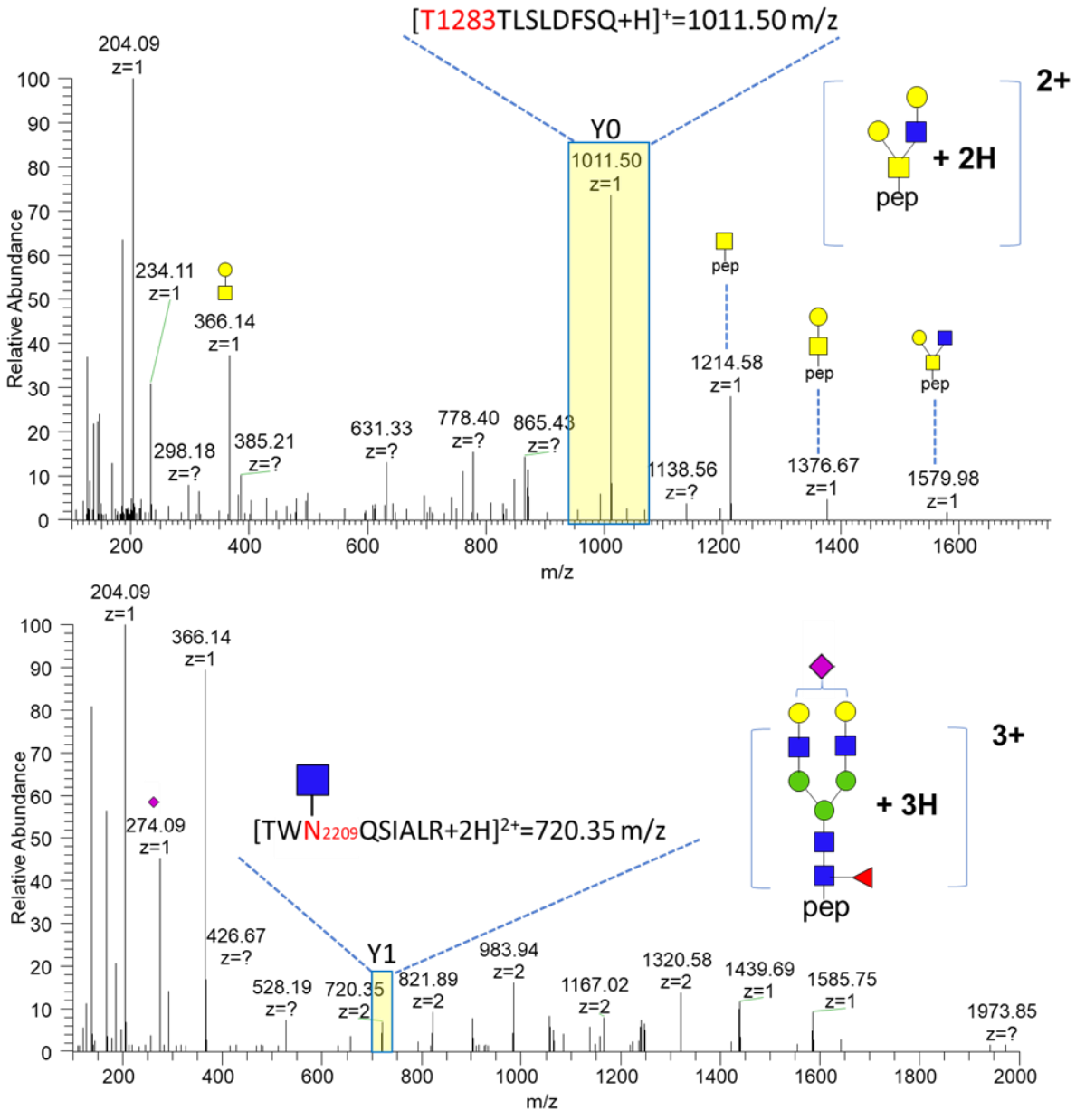


Figure 2.2 General HCD fragmentation spectra of N-glycopeptide and O-glycopeptide. Y0 and Y1 ions were highlighted.

Intact glycopeptides sequencing could achieve simultaneous elucidation of glycan structures and glycosites, which are fundamental to reveal biological functions of protein glycosylation.[37] Significant improvements have been made in characterizing intact glycopeptides using combined fragmentation including CID/ETD and HCD/ETD.[38, 39] However, combined fragmentation usually increases the complexity of MS/MS spectra and makes subsequence analysis more challenging. Although a few programs are available for automated analysis of MS/MS spectra of intact glycopeptides,[40] output results have been notoriously inaccurate due to the lack of a well-developed scoring algorithm[41] and glycoinformatics database.[42] In this study, we developed an analytical strategy for intact glycopeptide sequencing. All the spectra generated from *N*- and *O*-glycopeptides were selected and combined into an mgf file with an in-house software based on the presence of HexNAc fragment ions of *m/z* 126.0549, 138.0549, 144.0655, 168.0760 and 204.0866 from HCD spectra. Theoretical masses of all possible peptides were calculated by PeptideMass from ExPASy. Online software GlycoMod was further utilized for database searching, followed by manual verification of identified *N*-glycopeptides according to *N*-glycosites, characteristic peaks, and Y1 ions. Identified *O*-glycopeptides were verified manually according to the characteristic peaks and Y0 ions represented in HCD spectra. Meanwhile, to identify exact *O*-glycosites, associated ETD-MS/MS spectra was researched using pFind software, and the EXoO method was also used. Although key hemostatic proteins such as FVIII and VWF are widely known to be *O*-glycosylated, only a few *O*-glycosites and glycoforms were described using tandem-MS due to the complexity and the site-specific heterogeneity of the O-glycosylation.[29] By integrating the alternated-MS fragmentation strategy and improved data analysis, both *O*-glycosites and glycoforms can be profiled.

2.4.2 N-glycosite Microheterogeneity

Through combining multiple enzymatic digestions, HILIC enrichment of glycopeptides, and alternated HCD and ETD fragmentations, 40 site-specific N-glycoforms (20 are core-fucosylated) and 12 N-glycosites were identified from pdFV with the simultaneous determination of peptide sequences and glycoform compositions (Table 2.1). Most N-glycosites on asparagine (Asn) residues within consensus sequence (Asn-x-Thr/Ser/Cys) were observed with glycan microheterogeneity, with 8 N-glycosites are identified with three or more glycoforms. N-Glycopeptides assignment of HRMS results were listed in Table 2.3-2.7. N-glycosites indentification by ^{18}O labeled PNGase F digestion were listed in Table 2.21. Tandem MS annotations of N-glycopeptides were illustrated in Figure 2.7-2.11. Regardless of structural details either at the core or the non-reducing terminus of each glycoform, a total of 57 unique N-glycopeptides were identified. Among identified N-glycosites, our results showed that Asn382 is most heterogeneous with a total of 17 N-glycoforms, dominated by core-fucosylated ones. This site locates within the A2 region of the heavy chain of FV. Ten glycoforms, including complex and high-mannose types, were identified at Asn554. On domain B, 4 N-glycosites were identified with various glycoforms. Among these sites, Asn938 contains 7 glycoforms, among which 4 are core-fucosylated. Asn1074 and Asn1221 are occupied by complex and hybrid type N-glycans. The high percentage of core-fucosylated N-glycan in all identified glycoforms showed that pdFV is heavily glycosylated with core fucose. The core-fucosylation ratio in identified glycoforms is similar to that of FVIII.[23] The well-known immunogenic sialic acid form, N-glycolylneuraminic acid (Neu5Gc), was not observed in pdFV.

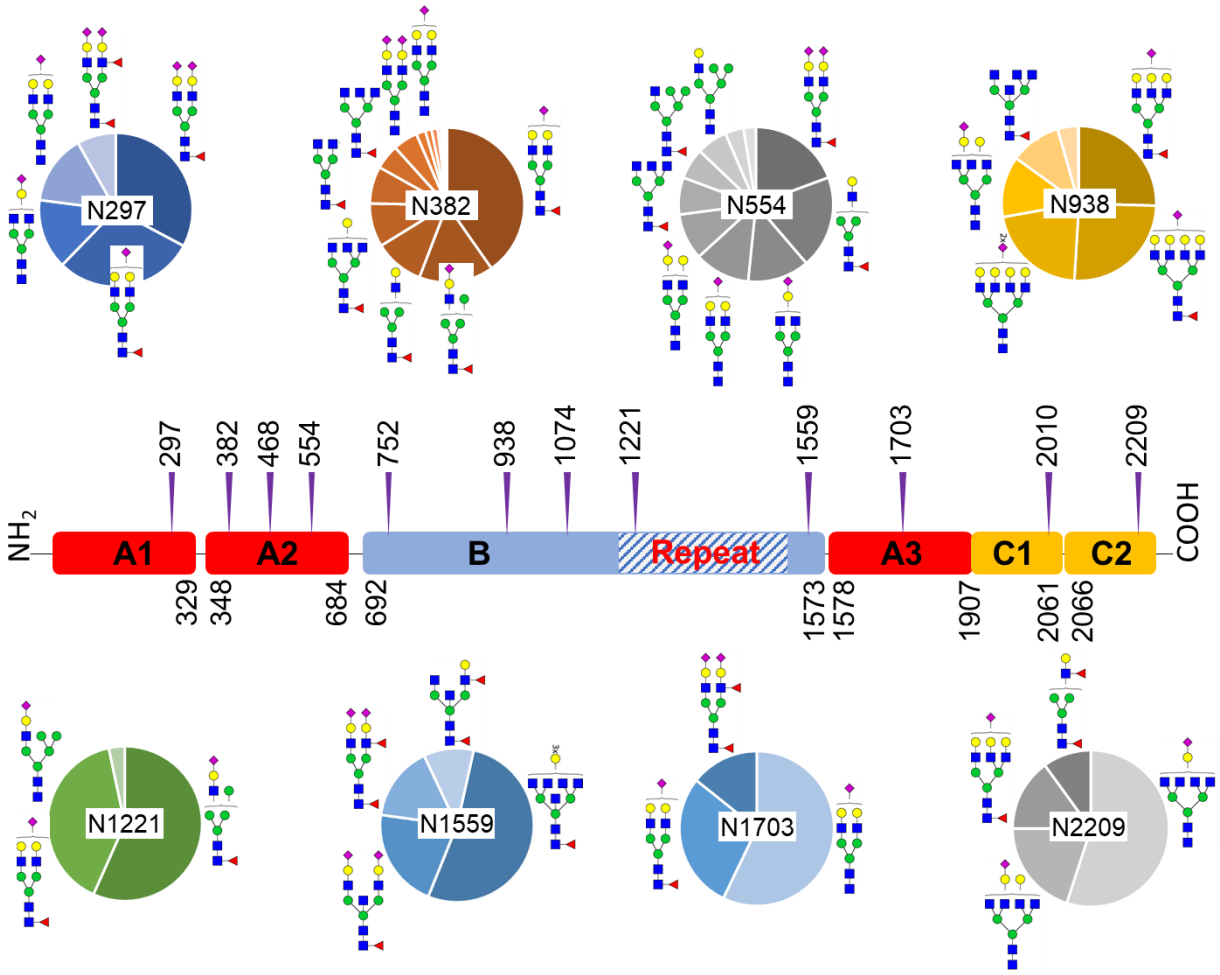


Figure 2.3 Microheterogeneity and relative abundance of N-glycoforms at 8 out of 12 FV N-glycosites.

2.4.3 Total relative abundances of N-glycoforms

We determined the relative abundance of detected glycoforms at each N-glycosite. Ion chromatograms of identified peptides extracted by LFQuant were used to quantify different glycopeptides. The mass-to-charge (m/z) and retention time (RT) of glycopeptide precursor ions identified above were used to extract ion chromatograms and calculate the peak area of individual glycopeptide. Peak areas were then normalized, and the relative abundance of each N-glycan composition was obtained regarding each site as shown in pie charts in Figure 2.3. Among 17 N-glycoforms identified at Asn382, monosialylated bi-antennary H5N4S1F1 (22), monosialylated

hybrid type H5N3S1F1 (11), and asialylated mono-antennary H4N3F1 (4) are highly abundant, with 40%, 16%, and 10%, respectively.

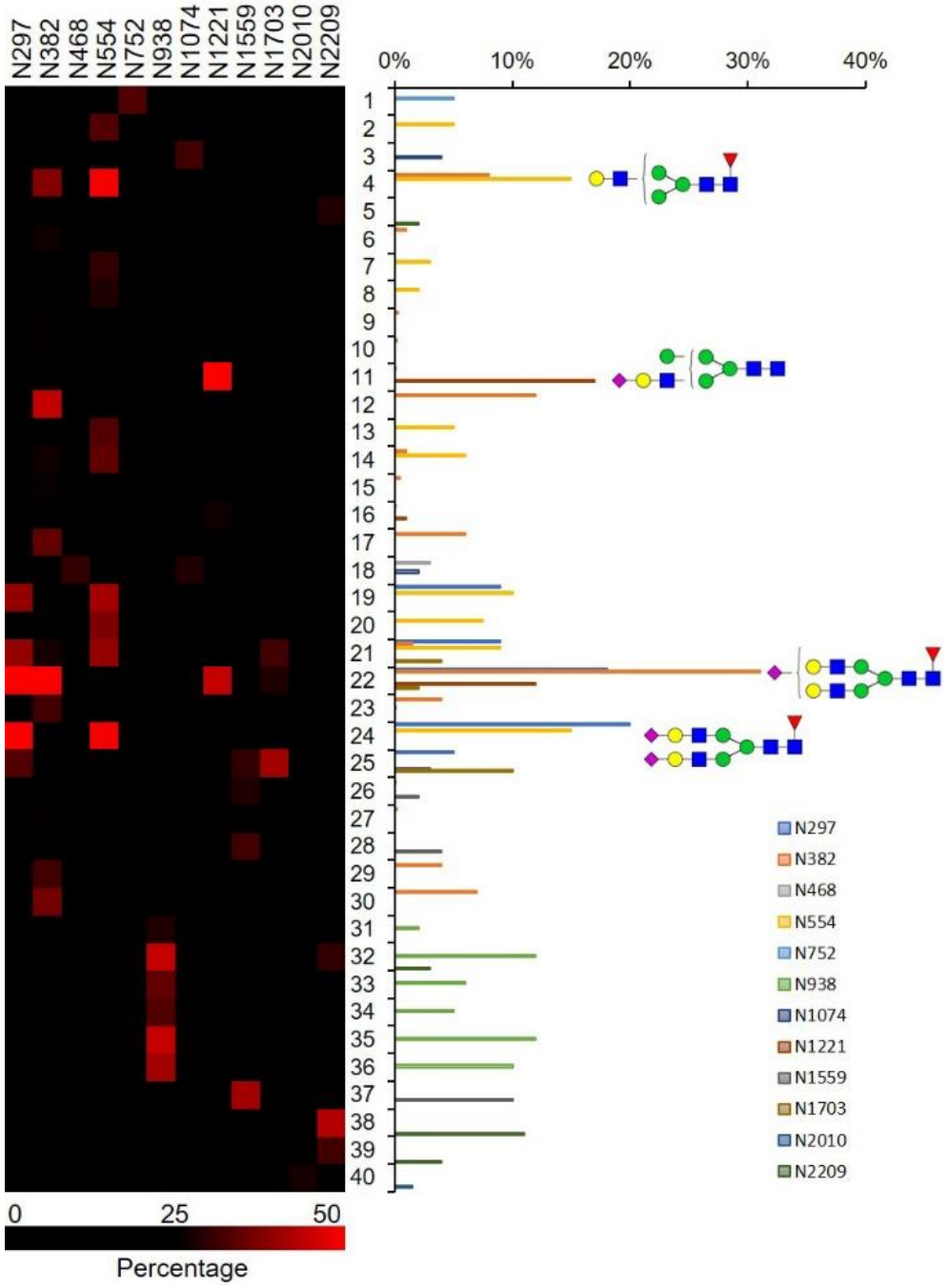


Figure 2.4 Distribution of N-glycan complexity. Total relative abundance of 40 N-glycans are represented in bar charts.

The top 3 highly abundant N-glycoform at Asn554 are all complex type, including disialylated bi-antennary H5N4S2F1 (24, 15%), asialyated H4N3F1 (4, 15%), and monosialylated bi-antennaryH4N4S1 (19, 13%). Figure 2.4 illustrates the general distribution of FV N-glycome based on the 40 N-glycoforms. Top 2 abundant N-glycans are all complex types. The most abundant structure is the monosialylated complex type H5N4S1F1 (22, 31.2%), which was defined on 4 N-glycosites, Asn297, Asn382, Asn1221, and Asn1703. The second abundant glycoform is the disialylated H5N4S2F1 (24, 20.1%) distributed on Asn297 and Asn382.

2.4.4 O-glycosite Microheterogeneity

FV was reported as highly O-glycosylated, with 53 possible sites (lacking glycoform information) identified from human serum samples, enriched by lectin VVA and PNA, which specific for Tn- and T-antigens.[43] All these O-glycosites are located in B domain, and many sites lack the exact position information.[43] We focused on identifying exact O-glycosites and glycoforms at each site. HCD-MS2 and ETD-MS2 were used to profile FV O-glycosyaltion, and Figure 2.5 illustrated corresponding spectra of one example glycopeptide APSHQQATT805AGSPLR from pdFV. O-Glycopeptides assignment of HRMS results were listed in Table 2.8-2.20. Tandem MS annotations of O-glycopeptides were shown in Figure 2.12-2.17. Additionally, a recently developed method EXoO was used identify and double confirm the O-glycosites. In this method, a unique enzyme, OpeRATOR, could selectively digest O-glycopeptide before the Thr/Ser reside where O-glycan locates, release a unique O-glycopeptide with an N-terminal Thr or Ser, thus enables the identification of the exact locations of O-glycosyaltion. The combination of the two methods enabled identification of totally 17 O-glycoforms on 26 O-glycosites (Table 2.2). As shown in Figure 2.6, 18 identified O-glycosites are localized within the B domain, 10 of which were identified before.[43] The B domain of FV contains

dozens of unusual 9 AA tandem repeat region,[44] the EXoO method was able to locate 4 O-glycopeptides in this region: T1211TLSPE, T1238TLSPD, T1247TLSLDLSQ and T1283TLSLDFSQ. Interestingly, O-glycoforms identified by HCD/ETD fragmentation on 3 sites (T1211, T1238, and T1283) share two sialylated structures (43, 45) attached to the Thr. We did not observe the other 43 possible O-glycosites identified before, which may be a result of different enrichment approaches.

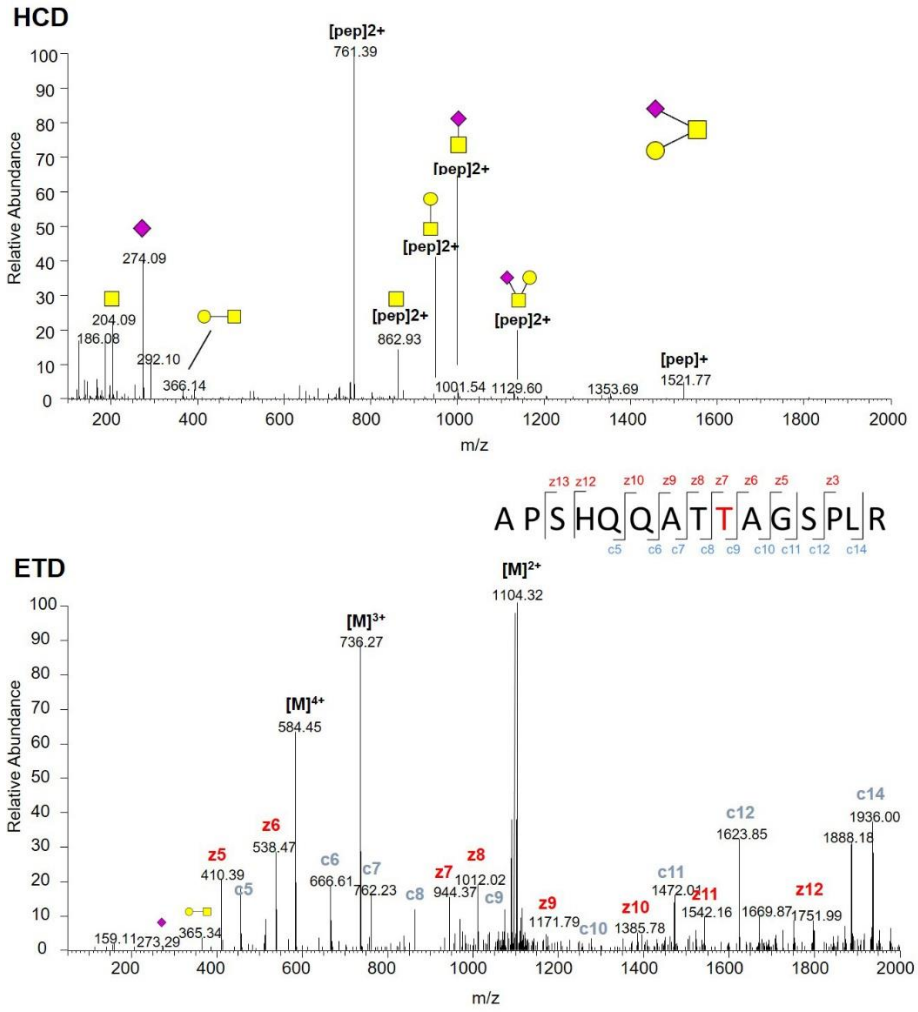


Figure 2.5 HCD-MS2 (above) and ETD-MS2 (bottom) spectra of (APSHQQATT805AG-SPLR) glycopeptide derived from plasma-derived FV.

The previous report applied lectin enrichment method is that more suitable for peptides with simple glycans such as Tn and T-antigens,[43] whereas this work used HILIC enrichment method based on hydrophilic interaction with hydroxyl group which is more suitable for peptides with larger complex glycans.

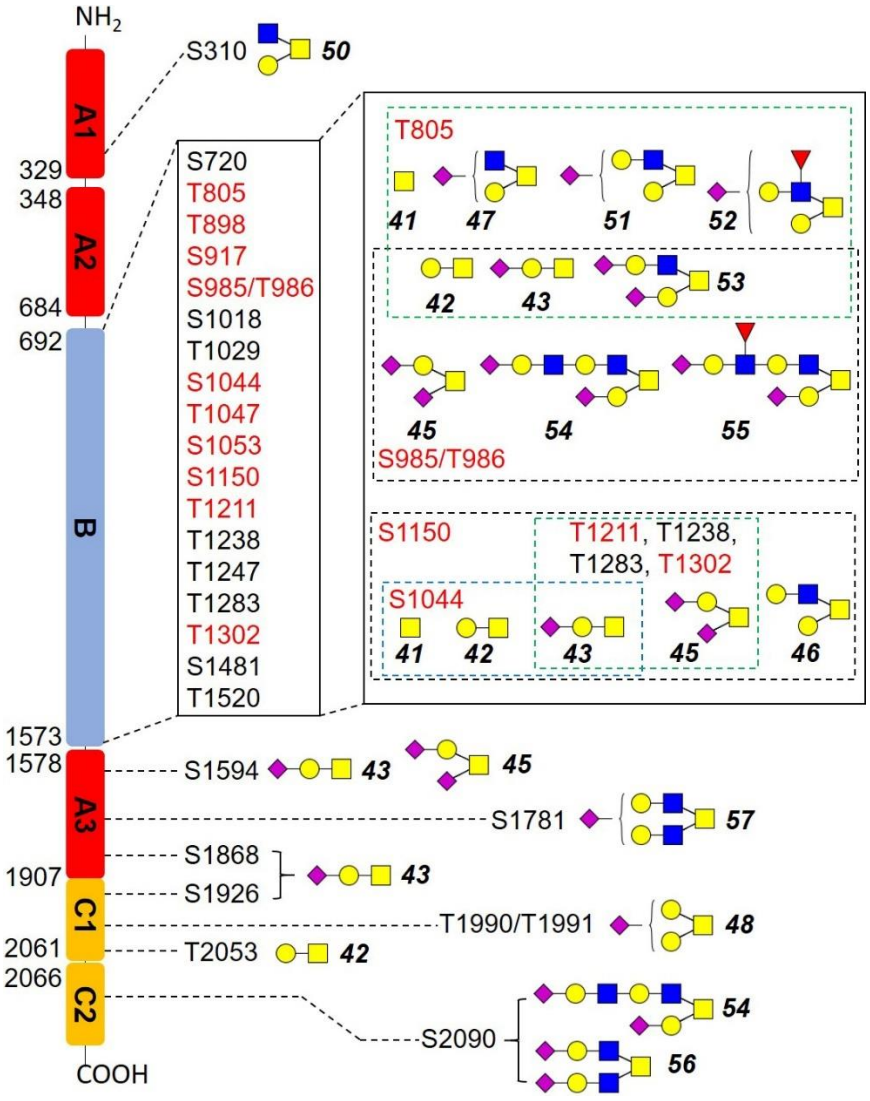


Figure 2.6 Identified O-glycosites and glycoforms on each site in this work. Red highlighted glycosites were also identified previously.

On the other hand, our method enabled the first identification of 1 O-glycosite on A1 domian (Ser310), 3 O-glycosites on A3 domian (Ser1594, Ser1781, and Ser1868), 3 O-glycosites on C1 domian (Ser1926, Thr1990/Thr1991, and Thr2053), and 1 O-glycosite on the C2 domain (Ser2090), most of which harbor complex O-glycans (Figure 2.6).

Among identified O-glycosites, Thr805 was most complex, with 7 O-glycoforms identified (Table 2.2). Ser985 or Thr986 were the second most complex O-glycosite with 6 glycoforms identified. The core 1 (42-45) and core 2 (46, 47, 49, 51-55) structures were the major O-glycoforms, while some core 4 O-glycoforms were also identified (56, 57). The predominant O-glycoforms were sialyl-T antigen (43), disialyl-T antigen (45) and T antigen (42). Worth noting that the O-glycans of pdFV is heavily sialylated. As shown in Table 2.2, N-acetylneuraminic acid (Neu5Ac) residues were observed on 13 of totally 17 identified O-glycoforms.

2.4.5 Biological significance of glycosylation study on FV

Site-specific relative abundance of glycopeptides is becoming increasingly important in protein medicine, especially for pharmaceutical quality control. We observed that FV is highly fucosylated and sialylated, with around 84.2% of detected N-glycans decorated with core Fucose (54.5%), Neu5Ac (50.9%), or both (22.8%), whereas non-core-fucosylated and non-sialylated glycoforms only account for about 15.8%. Such a high degree of sialylation is clearly important as whereas deglycosylated factor V activation was impaired.removal of sialic acid resulted in a 1.5-2-fold increase in clotting activity.[45] The von Willebrand factor (VWF) was also known to be highly sialylated hemostasis protein, which can bind to asialoglycoprotein receptor (ASGPR) in liver. Sialidase from infectious pathogens such as Streptococcus pneumoniae could descrease VWF sialylation thus increase its clearance.[46] Core-fucosyalted is closely reated with many

diseases, and it has been reported to regulate protein function. For example, depletion of core fucose on Asn-297 of IgG1 significantly increases antibody-dependent cellular cytotoxicity (ADCC) activity,[47] and deletion of core fucose on α3β1 Integrin reduces cell migration and attachment to extracellular matrix.[48] Given the high degree of core-fucosylation on FV, it may also play a functional role in clotting.

Nicolaes and co-workers reported that glycans at Asn2209 could impair the interaction between FVa and the phospholipid membrane.[14] Although glycosylation on Asn2209 is reported to affect FV function, glycoform information on this site is lacking. Our results revealed that Asn2209 is mainly occpupied by sialylated tetra-antennary N-glycan. We also identified another 11 N-glycosite on FV, 1 on A1 domain, 3 on A2 domain, 5 on B domain, 1 on A3 domain and another on C1 domain. It is reported that introduction of glycosylation on A2 and A3 domain of FV attenuated FXa binding,[17, 18] but the roles of its natural glycosylation on these domains remains unclear. Our results provide fundamentals to further studies regarding roles of such N-glycosylation on the function of FV. One potential site that of interest is Asn1703. Regardless of structural similarities between FV and FVIII, their N-glycosylation are not conserved except for Asn1703, which corresponding to Asn1810 in FVIII. Our result indicated that they are both occupied with complex type N-glycans. Asn1810 on FVIII is thought to be involved in the interaction with LRP and lipid, and consequently plays important roles in the clearance and clotting activity of FVIII, [49] whereas the role of Asn1703 on FV is unknow.

While O-glycosylation is thought to be on the B domain exclusilvely,[50, 51] we were able to identify 8 O-glycosites on A1, A3, C1 and C2 domains. Identified O-glycoforms of FV are mainly O-GalNAc core 1 structure, T (42), sialyl-T (43) and disialyl-T antigen (45) occupy most of the O-glycosites. Core 2 O-GalNAc structures were also identified on several O-glycosites especially

T805. Such mucin type O-GalNAc glycosylation is well known for its biological function in cancner and immune system, and altered expression of Tn (41), sTn (43), and T antigen (42) were usually highly expressed in different types of cancers including gastric, colon, breast, lung cancer.[52] However, O-glycosylation in FV and other heamostasis proteins had been rarely studied, and the potential roles of O-glycans are yet to be revealed. Additionally, it is reported that heterogeneity of glycoforms on glycoproteins affects protein stability, activity, and immunogenicity.[53, 54] We identified two heterogeneous O-glycosites (T805 with 7 glycoforms, S985/T986 with 6 glycoforms), and several N-glycosites with high microheterogeneity, e.g., Asn 382 (17 glycoforms) and Asn554 (10 glycoforms). These glycosites may be promising targets for further functional and mechanism study, and our results provide knowledge fundamental.

2.5 Conclusion

In summary, we comprehensively analyzed the glycosylation of pdFV by a strategy integrating multi-enzyme digestion, ZIC-HILIC enrichment, and alternated HCD and ETD. The implementation of different fragmentation strategies in one MS run and the combination of fragmentation strategies enabled us to acquire more information on both carbohydrate moieties and intact glycopeptides. Finally, a total of 57 unique N-glycopeptides and 51 O-glycopeptides were identified, categorized into 12 N-glycosites with 40 N-glycoforms, and 26 O-glycosites with 17 O-glycoforms, respectively. These glycosylation information is fundamental for further functional studies and FV-related therapeutic development.

2.6 References

1. Vehar, G. A.; Keyt, B.; Eaton, D.; Rodriguez, H.; O'Brien, D. P.; Rotblat, F.; Oppermann, H.; Keck, R.; Wood, W. I.; Harkins, R. N.; Tuddenham, E. G.; Lawn, R. M.; Capon, D. J., Structure of human factor VIII. *Nature* **1984,** *312* (5992), 337-42.
2. Kane, W. H.; Ichinose, A.; Hagen, F. S.; Davie, E. W., Cloning of cDNAs coding for the heavy chain region and connecting region of human factor V, a blood coagulation factor with four types of internal repeats. *Biochemistry* **1987,** *26* (20), 6508-14.
3. Huang, J. N.; Koerper, M. A., Factor V deficiency: a concise review. *Haemophilia : the official journal of the World Federation of Hemophilia* **2008,** *14* (6), 1164-9.
4. Kalafatis, M., Coagulation factor V: a plethora of anticoagulant molecules. *Current opinion in hematology* **2005,** *12* (2), 141-8.
5. Lee, W. S.; Chong, L. A.; Begum, S.; Abdullah, W. A.; Koh, M. T.; Lim, E. J., Factor V inhibitor in neonatal intracranial hemorrhage secondary to severe congenital factor V deficiency. *Journal of pediatric hematology/oncology* **2001,** *23* (4), 244-6.
6. Bulato, C.; Novembrino, C.; Anzoletti, M. B.; Spiezia, L.; Gavasso, S.; Berbenni, C.; Tagariello, G.; Farina, C.; Nardini, I.; Campello, E. J. H., "In vitro" correction of the severe factor V deficiency ‐ related coagulopathy by a novel plasma ‐ derived factor V concentrate. **2018,** *24* (4), 648-656.
7. Tracy, P. B.; Eide, L. L.; Bowie, E. J.; Mann, K. G., Radioimmunoassay of factor V in human plasma and platelets. *Blood* **1982,** *60* (1), 59-63.
8. Helenius, A., How N-linked oligosaccharides affect glycoprotein folding in the endoplasmic reticulum. *Molecular biology of the cell* **1994,** *5* (3), 253-65.
9. Mitra, N.; Sharon, N.; Surolia, A., Role of N-linked glycan in the unfolding pathway of Erythrina corallodendron lectin. *Biochemistry* **2003,** *42* (42), 12208-16.
10. Narhi, L. O.; Arakawa, T.; Aoki, K. H.; Elmore, R.; Rohde, M. F.; Boone, T.; Strickland, T. W., The effect of carbohydrate on the structure and stability of erythropoietin. *The Journal of biological chemistry* **1991,** *266* (34), 23022-6.
11. Shental-Bechor, D.; Levy, Y., Effect of glycosylation on protein folding: a close look at thermodynamic stabilization. *Proceedings of the National Academy of Sciences of the United States of America* **2008,** *105* (24), 8256-61.
12. Sharon, N.; Lis, H., Lectins as cell recognition molecules. *Science (New York, N.Y.)* **1989,** *246* (4927), 227-34.
13. Pittman, D. D.; Tomkinson, K. N.; Kaufman, R. J., Post-translational requirements for functional factor V and factor VIII secretion in mammalian cells. *The Journal of biological chemistry* **1994,** *269* (25), 17329-37.
14. Nicolaes, G. A.; Villoutreix, B. O.; Dahlback, B., Partial glycosylation of Asn2181 in human factor V as a cause of molecular and functional heterogeneity. Modulation of glycosylation efficiency by mutagenesis of the consensus sequence for N-linked glycosylation. *Biochemistry* **1999,** *38* (41), 13584-91.
15. Rosing, J.; Bakker, H. M.; Thomassen, M. C.; Hemker, H.; Tans, G. J. J. o. B. C., Characterization of two forms of human factor Va with different cofactor activities. **1993,** *268* (28), 21130-21136.
16. Silveira, J. R.; Kalafatis, M.; Tracy, P. B. J. B., Carbohydrate moieties on the procofactor factor V, but not the derived cofactor factor Va, regulate its inactivation by activated protein C. **2002,** *41* (5), 1672-1680.

17. Steen, M. r.; Norstrøm, E. A.; Tholander, A.-L.; Bolton-Maggs, P. H.; Mumford, A.; McVey, J. H.; Tuddenham, E. G.; Dahlbäck, B. r. J. B., Functional characterization of factor V-Ile359Thr: a novel mutation associated with thrombosis. **2004,** *103* (9), 3381-3387.
18. Steen, M.; Villoutreix, B. O.; Norstrøm, E. A.; Yamazaki, T.; Dahlbäck, B. J. J. o. B. C., Defining the factor Xa-binding site on factor Va by site-directed glycosylation. **2002,** *277* (51), 50022-50029.
19. Sodetz, J. M.; Paulson, J. C.; Pizzo, S. V.; McKee, P. A., Carbohydrate on human factor VIII/von Willebrand factor. Impairment of function by removal of specific galactose residues. *The Journal of biological chemistry* **1978,** *253* (20), 7202-6.
20. Kannicht, C.; Ramstrom, M.; Kohla, G.; Tiemeyer, M.; Casademunt, E.; Walter, O.; Sandberg, H., Characterisation of the post-translational modifications of a novel, human cell line-derived recombinant human factor VIII. *Thrombosis research* **2013,** *131* (1), 78-88.
21. Thim, L.; Vandahl, B.; Karlsson, J.; Klausen, N. K.; Pedersen, J.; Krogh, T. N.; Kjalke, M.; Petersen, J. M.; Johnsen, L. B.; Bolt, G.; Norby, P. L.; Steenstrup, T. D., Purification and characterization of a new recombinant factor VIII (N8). *Haemophilia : the official journal of the World Federation of Hemophilia* **2010,** *16* (2), 349-59.
22. Canis, K.; Anzengruber, J.; Garenaux, E.; Feichtinger, M.; Benamara, K.; Scheiflinger, F.; Savoy, L. A.; Reipert, B. M.; Malisauskas, M., In-depth comparison of N-glycosylation of human plasma-derived factor VIII and different recombinant products: from structure to clinical implications. *J Thromb Haemost* **2018**.
23. Kannicht, C.; Ramström, M.; Kohla, G.; Tiemeyer, M.; Casademunt, E.; Walter, O.; Sandberg, H. J. T. r., Characterisation of the post-translational modifications of a novel, human cell line-derived recombinant human factor VIII. **2013,** *131* (1), 78-88.
24. Lai, J. D.; Swystun, L. L.; Cartier, D.; Nesbitt, K.; Zhang, C.; Hough, C.; Dennis, J. W.; Lillicrap, D. J. h., N-linked glycosylation modulates the immunogenicity of recombinant human factor VIII in hemophilia A mice. **2018,** *103* (11), 1925-1936.
25. Singh, C.; Zampronio, C. G.; Creese, A. J.; Cooper, H. J. J. J. o. p. r., Higher energy collision dissociation (HCD) product ion-triggered electron transfer dissociation (ETD) mass spectrometry for the analysis of N-linked glycoproteins. **2012,** *11* (9), 4517-4525.
26. Küster, B.; Mann, M. J. A. c., 18O-labeling of N-glycosylation sites to improve the identification of gel-separated glycoproteins using peptide mass mapping and database searching. **1999,** *71* (7), 1431-1440.
27. Yang, W.; Ao, M.; Hu, Y.; Li, Q. K.; Zhang, H. J. M. s. b., Mapping the O‐glycoproteome using site‐specific extraction of O‐linked glycopeptides (EXoO). **2018,** *14* (11).
28. Wiśniewski, J. R.; Zougman, A.; Nagaraj, N.; Mann, M. J. N. m., Universal sample preparation method for proteome analysis. **2009,** *6* (5), 359.
29. Gashash, E. A.; Aloor, A.; Li, D.; Zhu, H.; Xu, X. Q.; Xiao, C.; Zhang, J.; Parameswaran, A.; Song, J.; Ma, C.; Xiao, W.; Wang, P. G., An Insight into Glyco-Microheterogeneity of Plasma von Willebrand Factor by Mass Spectrometry. *J Proteome Res* **2017,** *16* (9), 3348-3362.
30. Wilkins, M. R.; Lindskog, I.; Gasteiger, E.; Bairoch, A.; Sanchez, J. C.; Hochstrasser, D. F.; Appel, R. D., Detailed peptide characterization using PEPTIDEMASS--a World-Wide-Web-accessible tool. *Electrophoresis* **1997,** *18* (3-4), 403-8.
31. Ma, C.; Qu, J.; Li, X.; Zhao, X.; Li, L.; Xiao, C.; Edmunds, G.; Gashash, E.; Song, J.; Wang, P. G., Improvement of core-fucosylated glycoproteome coverage via alternating HCD and ETD fragmentation. *Journal of proteomics* **2016,** *146*, 90-8.

32. Shen, Y.; Tolic, N.; Purvine, S. O.; Smith, R. D., Improving collision induced dissociation (CID), high energy collision dissociation (HCD), and electron transfer dissociation (ETD) fourier transform MS/MS degradome-peptidome identifications using high accuracy mass information. *J Proteome Res* **2012,** *11* (2), 668-77.
33. Jia, W.; Lu, Z.; Fu, Y.; Wang, H. P.; Wang, L. H.; Chi, H.; Yuan, Z. F.; Zheng, Z. B.; Song, L. N.; Han, H. H.; Liang, Y. M.; Wang, J. L.; Cai, Y.; Zhang, Y. K.; Deng, Y. L.; Ying, W. T.; He, S. M.; Qian, X. H., A strategy for precise and large scale identification of core fucosylated glycoproteins. *Molecular & cellular proteomics : MCP* **2009,** *8* (5), 913-23.
34. Segu, Z. M.; Mechref, Y., Characterizing protein glycosylation sites through higher-energy C-trap dissociation. *Rapid communications in mass spectrometry : RCM* **2010,** *24* (9), 1217-25.
35. Thaysen-Andersen, M.; Wilkinson, B. L.; Payne, R. J.; Packer, N. H., Site-specific characterisation of densely O-glycosylated mucin-type peptides using electron transfer dissociation ESI-MS/MS. *Electrophoresis* **2011,** *32* (24), 3536-45.
36. Thaysen-Andersen, M.; Packer, N. H.; Schulz, B. L., Maturing Glycoproteomics Technologies Provide Unique Structural Insights into the N-glycoproteome and Its Regulation in Health and Disease. *Molecular & cellular proteomics : MCP* **2016,** *15* (6), 1773-90.
37. Cao, L.; Qu, Y.; Zhang, Z.; Wang, Z.; Prytkova, I.; Wu, S., Intact glycopeptide characterization using mass spectrometry. *Expert review of proteomics* **2016,** *13* (5), 513-22.
38. Scott, N. E.; Parker, B. L.; Connolly, A. M.; Paulech, J.; Edwards, A. V.; Crossett, B.; Falconer, L.; Kolarich, D.; Djordjevic, S. P.; Hojrup, P.; Packer, N. H.; Larsen, M. R.; Cordwell, S. J., Simultaneous glycan-peptide characterization using hydrophilic interaction chromatography and parallel fragmentation by CID, higher energy collisional dissociation, and electron transfer dissociation MS applied to the N-linked glycoproteome of Campylobacter jejuni. *Molecular & cellular proteomics : MCP* **2011,** *10* (2), M000031-mcp201.
39. Yin, X.; Bern, M.; Xing, Q.; Ho, J.; Viner, R.; Mayr, M., Glycoproteomic analysis of the secretome of human endothelial cells. *Molecular & cellular proteomics : MCP* **2013,** *12* (4), 956-78.
40. Hu, H.; Khatri, K.; Zaia, J., Algorithms and design strategies towards automated glycoproteomics analysis. *Mass spectrometry reviews* **2017,** *36* (4), 475-498.
41. Liu, G.; Cheng, K.; Lo, C. Y.; Li, J.; Qu, J.; Neelamegham, S., A comprehensive, open-source platform for mass spectrometry based glycoproteomics data analysis. *Molecular & cellular proteomics : MCP* **2017**.
42. Liu, G.; Neelamegham, S., Integration of systems glycobiology with bioinformatics toolboxes, glycoinformatics resources, and glycoproteomics data. *Wiley interdisciplinary reviews. Systems biology and medicine* **2015,** *7* (4), 163-81.
43. King, S. L.; Joshi, H. J.; Schjoldager, K. T.; Halim, A.; Madsen, T. D.; Dziegiel, M. H.; Woetmann, A.; Vakhrushev, S. Y.; Wandall, H. H. J. B. a., Characterizing the O-glycosylation landscape of human plasma, platelets, and endothelial cells. **2017,** *1* (7), 429-442.
44. Zhu, H.; Toso, R.; Camire, R. M., Inhibitory sequences within the B-domain stabilize circulating factor V in an inactive state. *The Journal of biological chemistry* **2007,** *282* (20), 15033-9.
45. BRUIN, T.; STURK, A.; TEN CATE, J. W.; CATH, M. J. E. j. o. b., The function of the human factor V carbohydrate moiety in blood coagulation. **1987,** *170* (1 - 2), 305-310.
46. Pipe, S. W.; Montgomery, R. R.; Pratt, K. P.; Lenting, P. J.; Lillicrap, D. J. B., The Journal of the American Society of Hematology, Life in the shadow of a dominant partner: the FVIII-VWF association and its clinical implications for hemophilia A. **2016,** *128* (16), 2007-2016.

47. Shields, R. L.; Lai, J.; Keck, R.; O'Connell, L. Y.; Hong, K.; Meng, Y. G.; Weikert, S. H.; Presta, L. G. J. J. o. B. C., Lack of fucose on human IgG1 N-linked oligosaccharide improves binding to human FcγRIII and antibody-dependent cellular toxicity. **2002,** *277* (30), 26733-26740.
48. Takahashi, M.; Kuroki, Y.; Ohtsubo, K.; Taniguchi, N. J. C. r., Core fucose and bisecting GlcNAc, the direct modifiers of the N-glycan core: their functions and target proteins. **2009,** *344* (12), 1387-1390.
49. Kosloski, M. P.; Miclea, R. D.; Balu-Iyer, S. V., Role of glycosylation in conformational stability, activity, macromolecular interaction and immunogenicity of recombinant human factor VIII. *The AAPS journal* **2009,** *11* (3), 424-31.
50. Kane, W. H.; Majerus, P. W. J. J. o. B. C., Purification and characterization of human coagulation factor V. **1981,** *256* (2), 1002-1007.
51. Fernández, J. A.; Hackeng, T. M.; Kojima, K.; Griffin, J. H. J. B., The Journal of the American Society of Hematology, The carbohydrate moiety of factor V modulates inactivation by activated protein C. **1997,** *89* (12), 4348-4354.
52. Fu, C.; Zhao, H.; Wang, Y.; Cai, H.; Xiao, Y.; Zeng, Y.; Chen, H. J. H., Tumor‐associated antigens: Tn antigen, sTn antigen, and T antigen. **2016,** *88* (6), 275-286.
53. Kolarich, D.; Lepenies, B.; Seeberger, P. H. J. C. o. i. c. b., Glycomics, glycoproteomics and the immune system. **2012,** *16* (1-2), 214-220.
54. Takahashi, M.; Tsuda, T.; Ikeda, Y.; Honke, K.; Taniguchi, N. J. G. j., Role of N-glycans in growth factor signaling. **2003,** *20* (3), 207-212.

2.7 Appendix

Table 2.1 List of 40 identified N-glycoforms and 12 N-glycosites from FV.

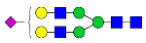
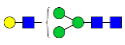
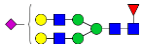
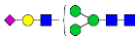
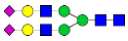
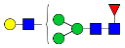
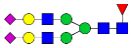
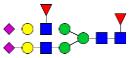
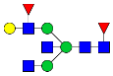
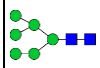
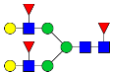
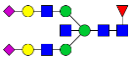
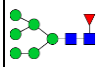
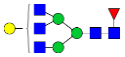
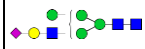
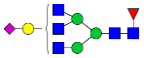
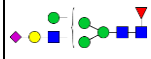
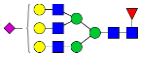
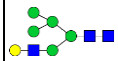
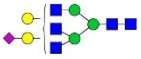
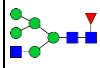
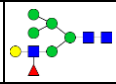
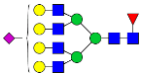
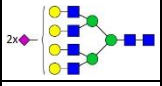
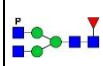
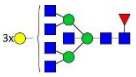
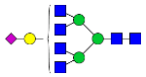
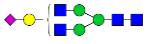
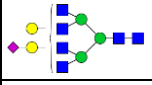
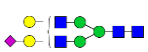
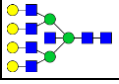
No.	Glycoform	Glycosite		No.	Glycoform	Glycosite
1		752		**21**		297, 382, 554, 1703
2		554		**22**		297, 382, 1221,1703
3		1074		**23**		382, 938
4		382, 554		**24**		297, 554
5		2209		**25**		297, 1559, 1703
6		382		**26**		382,1559
7		554		**27**		382
8		554		**28**		1559
9		382		**29**		382
10		382		**30**		382
11		382, 1221		**31**		938
12		382		**32**		938, 2209
13		554		**33**		938
14		382, 554		**34**		938
15		382		**35**		938
16		382, 1221		**36**		938
17		382		**37**		1559
18		468, 1074		**38**		2209
19		297, 554		**39**		2209
20		554		**40**		2010

Table 2.2 List of 17 identified O-glycoforms and 26 O-glycosites from FV.

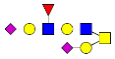
No.	Glycoform	Glycosite			No.	Glycoform	Glycosite
41		T805, S1044, S1150			**50**		S720
42		T805, S985/T986, S1044, T1047, S1053, S1150, S1926, T2053			**51**		T805, S1018
43		T805, T898, S917 S985/T986, S1044, T1047, S1150, T1211, T1238, T1283, T1302, S1481, T1520, S1594, S1868, S1926			**52**		T805, T1029
44		S917, S1053			**53**		T805, S985/T986
45		S985/T986, S1053, S1150, T1211, T1238, T1283, T1302, S1594			**54**		S985/T986, S2090
46		S1150, T1247			**55**		S985/T986
47		T805			**56**		S2090
48		T1990/T1991			**57**		S1781
49		S310					

Table 2.3 Identified N-glycopeptides in human derived FV N554: N554KSWYLE and N554KSWYLED

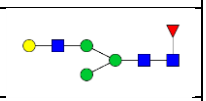
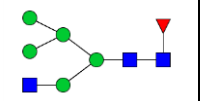
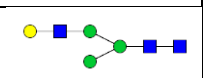
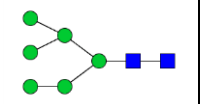
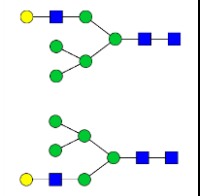
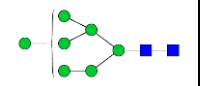
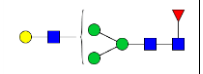
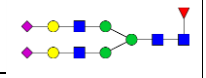
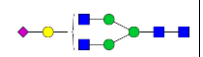
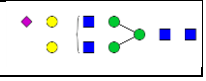
Glycan mass	Δmass (ppm)	Structure	Type	Peptide mass	Glycopeptide mass	CFG structure
1403.507	8.067	(Hex)1(HexNAc)1(Deoxyhexose)1+ (Man)3(GlcNAc)2	hybrid/complex	1053.477	2457.991	
1565.56	3.714	(Hex)2(HexNAc)1(Deoxyhexose)1+ (Man)3(GlcNAc)2	hybrid/complex	1053.477	2620.044	
1257.449	7.366	(Hex)1(HexNAc)1+ (Man)3(GlcNAc)2	hybrid/complex	1053.477	2311.933	
1378.476	7.575	(Hex)3+(Man)3(GlcNAc)2	high_man	1053.477	2432.960	
1581.555	6.878	(Hex)3(HexNAc)1+ (Man)3(GlcNAc)2	hybrid/complex	1053.477	2636.039	
1540.528	7.95	(Hex)4+(Man)3(GlcNAc)2	high_man	1053.477	2595.012	
1403.507	8.067	(Hex)1(HexNAc)1(Deoxyhexose)1+ (Man)3(GlcNAc)2	hybrid/complex	1053.477	2457.991	
2432.884	-1.829	(HexNAc)4(Deoxyhexose)1(NeuAc)2+(Man)3(GlcNAc)2	hybrid/complex	938.45	3372.341	
1751.624	16.463	(Hex)1(HexNAc)2(NeuAc)1 + (Man)3(GlcNAc)2	hybrid/complex	2315.111	4067.742	
1954.704	16.061	(Hex)1(HexNAc)3(NeuAc)1 + (Man)3(GlcNAc)2	hybrid/complex	2331.106	4286.817	

Table 2.4 Identified N-glycopeptides in human derived FV N1074: RRLKHSLVLHKSN1074E

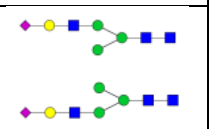
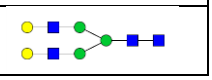
Glycan mass	Δmass (ppm)	Structure	Type	Peptide mass	Glycopeptide mass	CFG structure
1548.545	-3.083	(Hex)1(HexNAc)1(NeuAc)1+ (Hex)3(HexNAc)2	hybrid/complex	1715.991	3265.543	
1622.582	0.210	(Hex)2(HexNAc)2+ (Hex)3(HexNAc)2	hybrid/complex	1715.991	3339.58	

Table 2.5 Identified N-glycopeptides in human derived FV N382: SQHLDN382FSNQIGK and SQHLDN382FSNQIGKHYK

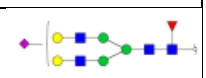
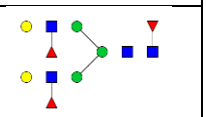
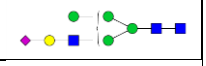
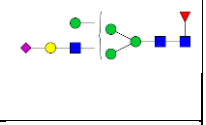
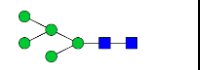
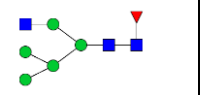
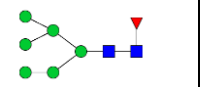
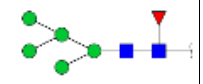
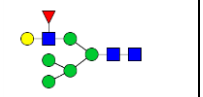
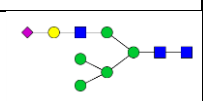
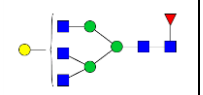
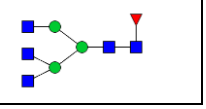
Glycan mass	Δmass (ppm)	Structure	Type	Peptide mass	Glycopeptide mass	CFG structure
2059.735	-6.032	(Hex)2(HexNAc)2(Deoxyhexose)1 (NeuAc)1+(Man)3(GlcNAc)2	hybrid/complex	1486.716	3547.459	
2060.755	-17.238	(Hex)2(HexNAc)2(Deoxyhexose)3+ (Man)3(GlcNAc)2	hybrid/complex	1486.716	3548.478	
1710.598	-8.994	(Hex)2(HexNAc)1(NeuAc)1+ (Man)3(GlcNAc)2	hybrid/complex	1486.716	3198.321	
1856.656	-5.462	(Hex)2 (HexNAc)1(Deoxyhexose)1 (NeuAc)1+(Man)3(GlcNAc)2	hybrid/complex	1486.716	3344.379	
1913.677	-6.252	(Hex)2(HexNAc)2(NeuAc)1+(Man)3(GlcNAc)2	hybrid/complex	1486.716	3401.400	
2204.772	2.473	(Hex)2(HexNAc)2(NeuAc)2+ (Man)3(GlcNAc)2	hybrid/complex	1486.716	3692.495	
1216.423	-9.289	(Hex)2+(Man)3(GlcNAc)2	high_man	1914.934	3132.364	
1565.560	-3.963	(Hex)2(HexNAc)1(Deoxyhexose)1+ (Man)3(GlcNAc)2	hybrid/complex	1914.934	3481.501	
1403.507	-4.668	(Hex)1(HexNAc)1(Deoxyhexose)1+ (Man)3(GlcNAc)2	hybrid/complex	1914.934	3319.448	
1524.534	-5.027	(Hex)3(Deoxyhexose)1+ (Man)3(GlcNAc)2	high_man	1914.934	3440.475	
1362.481	-8.012	(Hex)2(Deoxyhexose)1+(Man)3(GlcNAc)2	high_man	1914.934	3278.422	
1524.500	4.398	(HexNAc)2(Deoxyhexose)1(Phos)1+(Man)3(GlcNAc)2	hybrid/complex	1914.934	3440.441	
1727.613	-4.251	(Hex)3 (HexNAc)1(Deoxyhexose)1+ (Man)3(GlcNAc)2	hybrid/complex	1486.716	3215.336	
2018.708	-8.746	(Hex)3 (HexNAc)1 (Deoxyhexose)1 (NeuAc)1 + (Man)3(GlcNAc)2	hybrid/complex	1486.716	3506.431	
1809.666	5.905	(Hex)1(HexNAc)3(Deoxyhexose)1 +(Man)3(GlcNAc)2	hybrid/complex	1914.934	3725.607	
1647.613	5.556	(HexNAc)3(Deoxyhexose)1+(Man)3(GlcNAc)2	hybrid/complex	1914.934	3563.554	

Table 2.6 Identified N-glycopeptides in human derived FV N938: QSN938SSK, QSN938SSKILVGRWHLASEK, and LLLLKQSN938SSKILVGRWHLASE

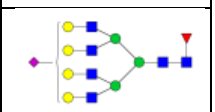
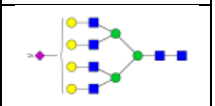
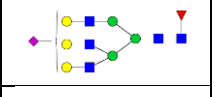
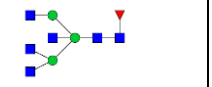
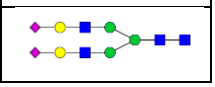
Glycan mass	Δmass (ppm)	Structure	Type	Peptide mass	Glycopeptide mass	CFG structure
2789.999	26.023	(Hex)4(HexNAc)4(Deoxyhexose)1 (NeuAc)1+(Man)3(GlcNAc)2	hybrid/complex	649.303	3440.309	
2935.037	9.129	(Hex)4(HexNAc)4(NeuAc)2+(Man)3(GlcNAc)2	hybrid/complex	649.303	3585.347	
2715.963	15.337	(Hex)3 (HexNAc)3 (Deoxyhexose)1 (NeuAc)2+(Man)3(GlcNAc)2	hybrid/complex	649.303	3366.273	
1938.709	11.046	(Hex)1(HexNAc)3(Deoxyhexose)1(NeuAc)1+(Man)3(GlcNAc)2	hybrid/complex	2039.091	3978.807	
1850.693	10.006	(HexNAc)4(Deoxyhexose)1+(Man)3(GlcNAc)2	hybrid/complex	2039.091	3890.791	
2204.772	3.178	(Hex)2(HexNAc)2(NeuAc)2+ (Man)3(GlcNAc)2	hybrid/complex	2491.427 (Glu-C)	4697.206	

Table 2.7 Identified N-glycopeptides in human derived FV N2209: TWN2209QSIALR or FIRVIPKTWN2209QSIALR

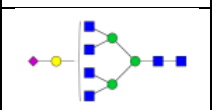
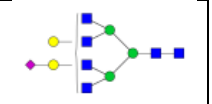
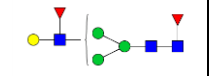
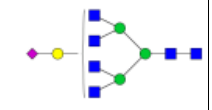
Glycan mass	Δmass (ppm)	Structure	Type	Peptide mass	Glycopeptide mass	CFG structure
2157.783	14.764	(Hex)1(HexNAc)4(NeuAc)1+ (Man)3(GlcNAc)2	hybrid/complex	1087.577	3246.367	
2319.836	15.852	(Hex)2(HexNAc)4(NeuAc)1+ (Man)3(GlcNAc)2	hybrid/complex	1087.577	3408.42	
2715.963	26.642	(Hex)3(HexNAc)3(Deoxyhexose)1 (NeuAc)2+(Man)3(GlcNAc)2	hybrid/complex	1957.126	4674.096	
1590.592	-40.97	(HexNAc)2(Deoxyhexose)2+(M an)3(GlcNAc)2	hybrid/complex	1524.878	3116.477	
2157.783	-27.589	(Hex)1(HexNAc)4(NeuAc)1+(Man)3(GlcNAc)2	hybrid/complex	1540.873	3699.663	
1913.677	1.852	(Hex)2(HexNAc)2(NeuAc)1+ (Man)3(GlcNAc)2	hybrid/complex	1524.878	3439.562	

Table 2.8 Identified O-glycopeptides in human derived FV S1053: S1053EAYNTFSER

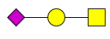
Glycan mass	Δmass (ppm)	Structure	Type	Peptide mass [M]	Glycopeptide mass	CFG structure
656.228	2.672	(Hex)1(HexNAc)1(NeuAc)1	Trypsin	1202.52	1859.756	
365.132	2.531	(Hex)1(HexNAc)1	Trypsin	1202.52	1568.66	

Table 2.9 Identified O-glycopeptides in human derived FV S1594: YS1594EFVQRE

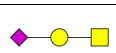
Glycan mass	Δmass (ppm)	Structure	Type	Peptide mass [M]	Glycopeptide mass	CFG structure
656.228	1.943	(Hex)1(HexNAc)1(NeuAc)1	Glu-C	1056.488	1713.723	
947.323	-2.029	(Hex)1(HexNAc)1(NeuAc)2	Glu-C	1056.488	2004.818	

Table 2.10 Identified O-glycopeptides in human derived FV 985,986 AWGESTPLANKPGK

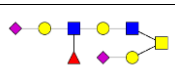
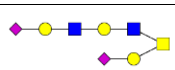
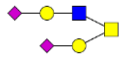
Glycan mass	Δmass (ppm)	Structure	Type	Peptide mass [M]	Glycopeptide mass	CFG structure
1823.645	-11.764	(Hex)3(HexNAc)3(Deoxyhexose)1 (NeuAc)2	Trypsin	1470.747	3295.399	
1677.587	-12.691	(Hex)3(HexNAc)3(NeuAc)2	Trypsin	1470.747	3149.341	
1312.455	-17.947	(Hex)2(HexNAc)2(NeuAc)2	Trypsin	1470.747	2784.209	
947.323	-3.286	(Hex)1(HexNAc)1(NeuAc)2	Trypsin	1454.752	2403.082	
656.228	-3.929	(Hex)1(HexNAc)1(NeuAc)1	Trypsin	1454.752	2111.987	
365.132	-1.592	(Hex)1(HexNAc)1	Trypsin	1454.752	1820.891	

Table 2.11 Identified O-glycopeptides in human derived FV T1047: T1047FHPLR

Glycan mass	Δmass (ppm)	Structure	Type	Peptide mass [M]	Glycopeptide mass	CFG structure
365.132	0.396	(Hex)1(HexNAc)1	Trypsin	769.423	1135.563	
656.228	3.035	(Hex)1(HexNAc)1(NeuAc)1	Trypsin	769.423	1426.658	

Table 2.12 Identified O-glycopeptides in human derived FV S1926: AS1926EFLGYWEPR

Glycan mass	Δmass (ppm)	Structure	Type	Peptide mass [M]	Glycopeptide mass	CFG structure
656.228	-0.951	(Hex)1 (HexNAc)1 (NeuAc)1	Trypsin	1353.635	2010.871	
365.132	-0.951	(Hex)1(HexNAc)1	Trypsin	1353.635	1719.775	

Table 2.13 Identified O-glycopeptides in human derived FV S1150 SSS1150PELSEMLEYDR

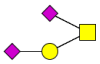
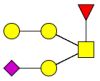
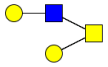
Glycan mass	Δmass (ppm)	Structure	Type	Peptide mass [M]	Glycopeptide mass	CFG structure
947.323	0.131	(Hex)1(HexNAc)1(NeuAc)2	Trypsin	1641.719	2590.049	
656.228	2.497	(Hex)1(HexNAc)1(NeuAc)1	Trypsin	1641.719	2298.954	
203.079	1.484	HexNAc	Trypsin	1641.719	1845.805	
1126.391	15.326	(Hex)3(HexNAc)1(Deoxyhexose)1 (NeuAc)1	Trypsin	1641.719	2769.117	
365.132	1.663	(Hex)1(HexNAc)1	Trypsin	1641.719	2007.858	
730.264	3.692	(Hex)2(HexNAc)2	Trypsin	1657.714	2388.985	

Table 2.14 Identified O-glycopeptides in human derived FV T805 APSHQQATT805AGSPLR

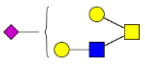
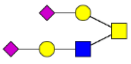
Glycan mass	Δmass (ppm)	Structure	Type	Peptide mass [M]	Glycopeptide mass	CFG structure
1021.36	-6.011	(Hex)2(HexNAc)2(NeuAc)1	Trypsin	1520.77	2543.137	
859.307	-3.145	(Hex)1(HexNAc)2(NeuAc)1	Trypsin	1520.77	2381.084	
1167.418	-0.248	(Hex)2(HexNAc)2 (Deoxyhexose)1(NeuAc)1	Trypsin	1520.77	2689.195	
203.079	-2.196	(HexNAc)1	Trypsin	1520.77	1724.856	
365.132	1.436	(Hex)1(HexNAc)1	Trypsin	1520.77	1886.909	
656.228	-3.300	(Hex)1(HexNAc)1(NeuAc)1	Trypsin	1520.77	2178.005	
1021.36	-1.017	(Hex)2(HexNAc)2(NeuAc)1	Trypsin	1520.77	2543.137	
1312.455	2.657	(Hex)2(HexNAc)2(NeuAc)2	Trypsin	1520.77	2834.232	

Table 2.15 Identified O-glycopeptides in human derived FV S1044: HTHHAPLS1044PR

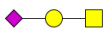
Glycan mass	Δmass (ppm)	Structure	Type	Peptide mass [M]	Glycopeptide mass	CFG structure
656.228	-0.751	(Hex)1(HexNAc)1(NeuAc)1	Trypsin	1151.595	1808.83	
365.132	-1.686	(Hex)1(HexNAc)1	Trypsin	1151.595	1517.727	
203.079	1.088	(HexNAc)1	Trypsin	1151.595	1354.674	

Table 2.16 Identified O-glycopeptides in human derived FV S1053: KHTHHAPLSPRTFH-PLRS1053E

Glycan mass	Δmass (ppm)	Structure	Type	Peptide mass [M]	Glycopeptide mass	CFG structure
656.228	-0.712	(Hex)1(HexNAc)1(NeuAc)1	Glu-C	2247.177	2904.413	

Table 2.17 Identified O-glycopeptides in human derived FV T1211: LSHT1211TLSPE

Glycan mass	Δmass (ppm)	Structure	Type	Peptide mass [M]	Glycopeptide mass	CFG structure
656.228	0.004	(Hex)1(HexNAc)1(NeuAc)1	Glu-C	983.492	1640.727	
947.323	2.707	(Hex)1(HexNAc)1(NeuAc)2	Glu-C	983.492	1931.822	

Table 2.18 Identified O-glycopeptides in human derived FV 1298-1306 LSHMTLSPE

Glycan mass	Δmass (ppm)	Structure	Type	Peptide mass [M]	Glycopeptide mass	CFG structure
656.228	1.897	(Hex)1(HexNAc)1(NeuAc)1	Glu-C	1013.485	1670.72	
947.323	2.737	(Hex)1(HexNAc)1(NeuAc)2	Glu-C	1013.485	1961.815	

Table 2.19 Identified O-glycopeptides in human derived FV T1238 LSHT1238TLSPD

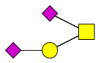
Glycan mass	Δmass (ppm)	Structure	Type	Peptide mass [M]	Glycopeptide mass	CFG structure
947.323	1.788	(Hex)1(HexNAc)1(NeuAc)2	Glu-C	969.477	1917.807	

Table 2.20 Identified O-glycopeptides in human derived FV T1283 LSHT1283TLSLD

Glycan mass	Δmass (ppm)	Structure	Type	Peptide mass [M]	Glycopeptide mass	CFG structure
947.323	1.619	(Hex)1(HexNAc)1(NeuAc)2	Glu-C	985.508	1933.838	

Table 2.21 N-glycosylation sites of human derived Factor V (Trypsin and Glu-C)

Peptides	**Modification**	**Score**	**MW**	**Sites**
VSAITLVSATSTTAnMTVGPEGK	15,N(18O_deamidated); 16,M(Oxidation[M]);	2.18E-26	2253.12955	297
NMASRPYSIYPHGVTFSPYEDEVnSSFTSGR	24,N(18O_deamidated);	3.20E-11	3497.57192	468
PYSIYPHGVTFSPYEDEVnSSFTSGR	19,N(18O_deamidated);	7.74E-27	2938.31826	468
GSYEIIQDTDEDTAVNNWLISPQnASR	24,N(18O_deamidated);	7.23E-26	3038.39905	977
TNInSSRDPDNIAAWYLR	4,N(18O_deamidated);	1.63E-11	2108.01724	1559
TYEDDSPEWFKEDNAVQPnSSYTYVWHATER	14,N(18O_deamidated); 19,N(18O_deamidated);	2.47E-15	3769.61036	1703
EDNAVQPnSSYTYVWHATER	8,N(18O_deamidated);	1.99E-23	2369.04457	1703
TWnQSIALRLELFGCDIY	15,C(Carbamidomethyl[C]);	4.99E-19	2198.08299	2209
TWnQSIALR	3,N(18O_deamidated);	2.49E-15	1090.56569	2209

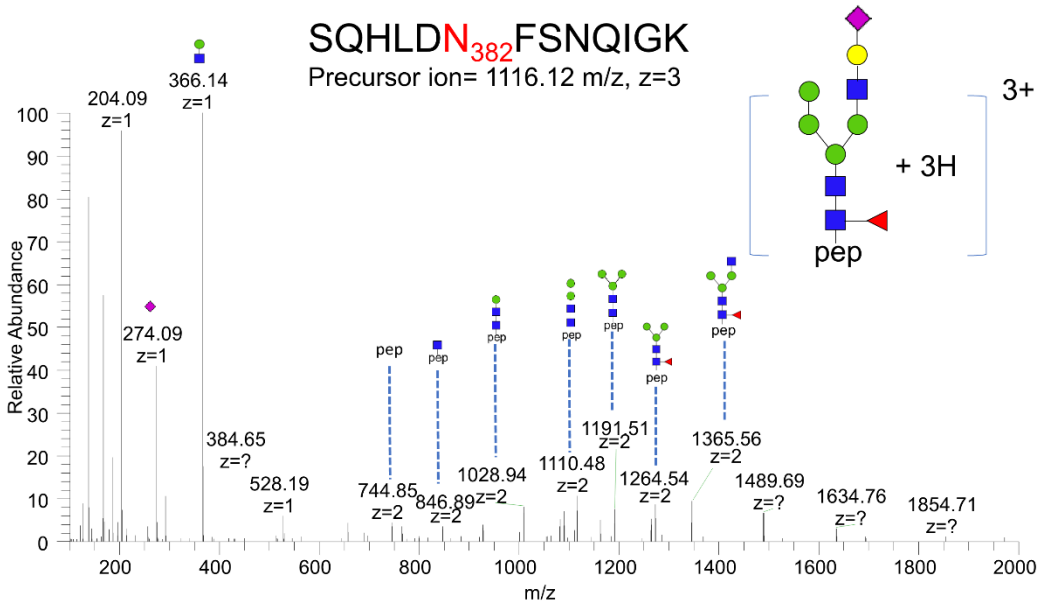


Figure 2.7 Tandem MS annotation of N-glycopeptide SQHLDN382FSNQIGK with hybrid type glycoform.

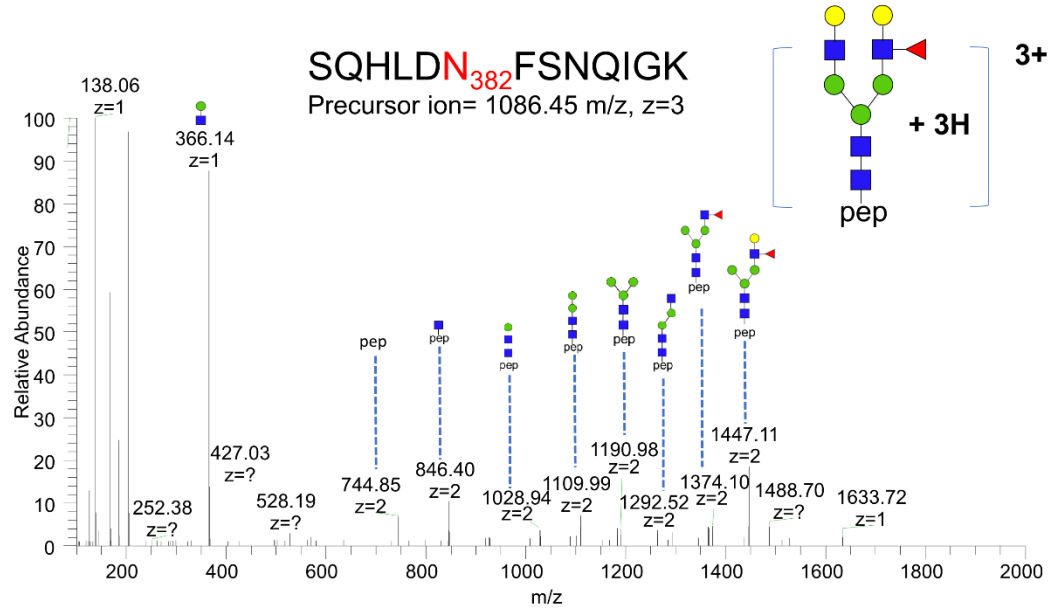


Figure 2.8 Tandem MS annotation of N-glycopeptide SQHLDN382FSNQIGK with complex type glycoform.

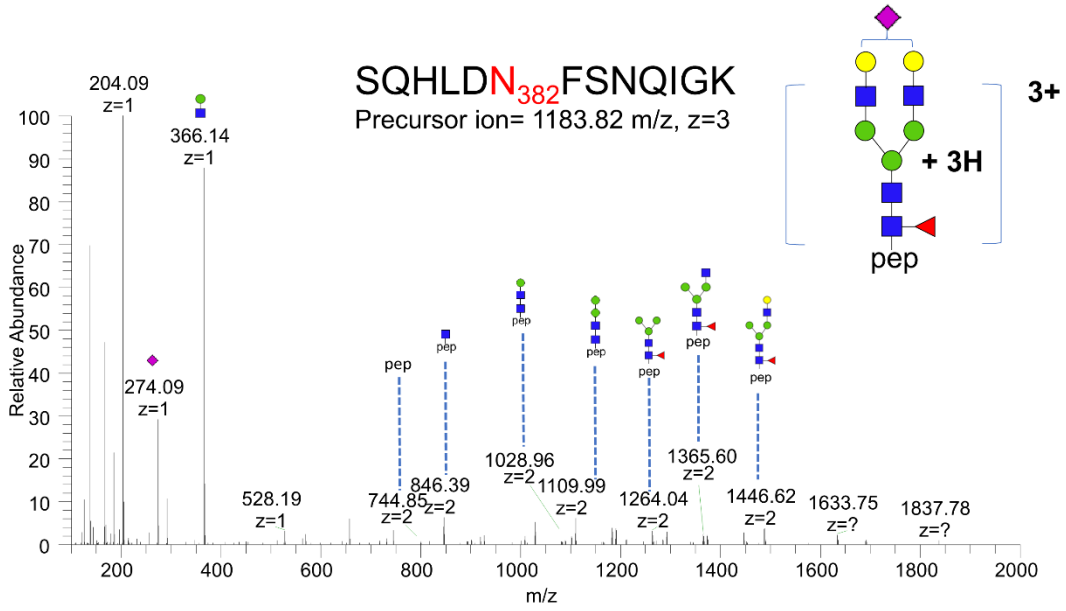


Figure 2.9 Tandem MS annotation of N-glycopeptide SQHLDN382FSNQIGK with sialylated complex type glycoform.

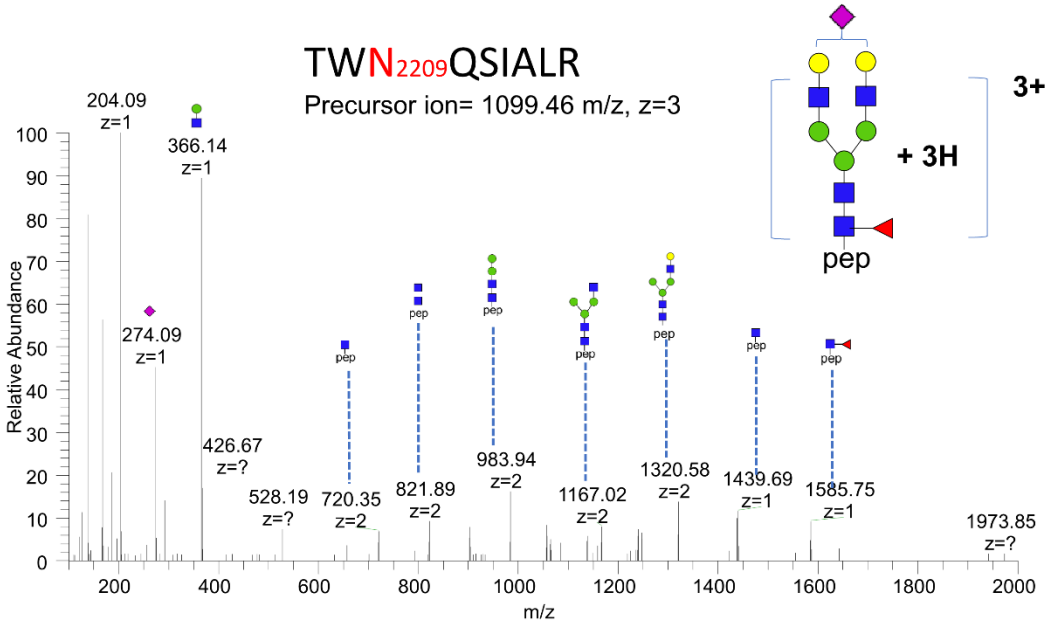


Figure 2.10 Tandem MS annotation of N-glycopeptide TWN382QSIALR with sialylated complex type glycoform.

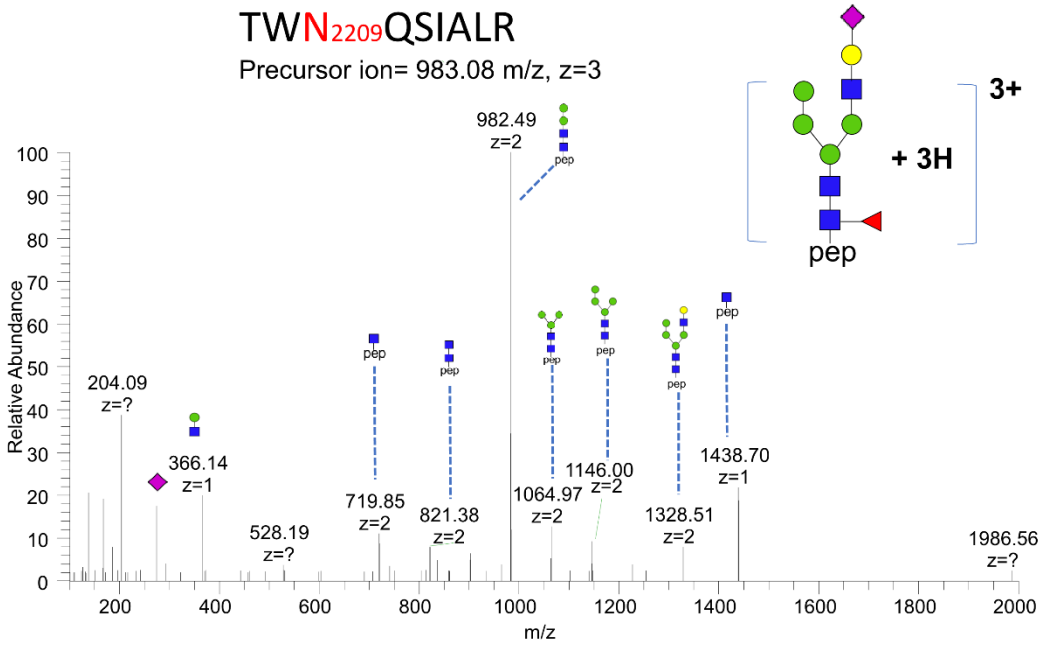


Figure 2.11 Tandem MS annotation of N-glycopeptide TWN382QSIALR with sialylated hybrid type glycoform.

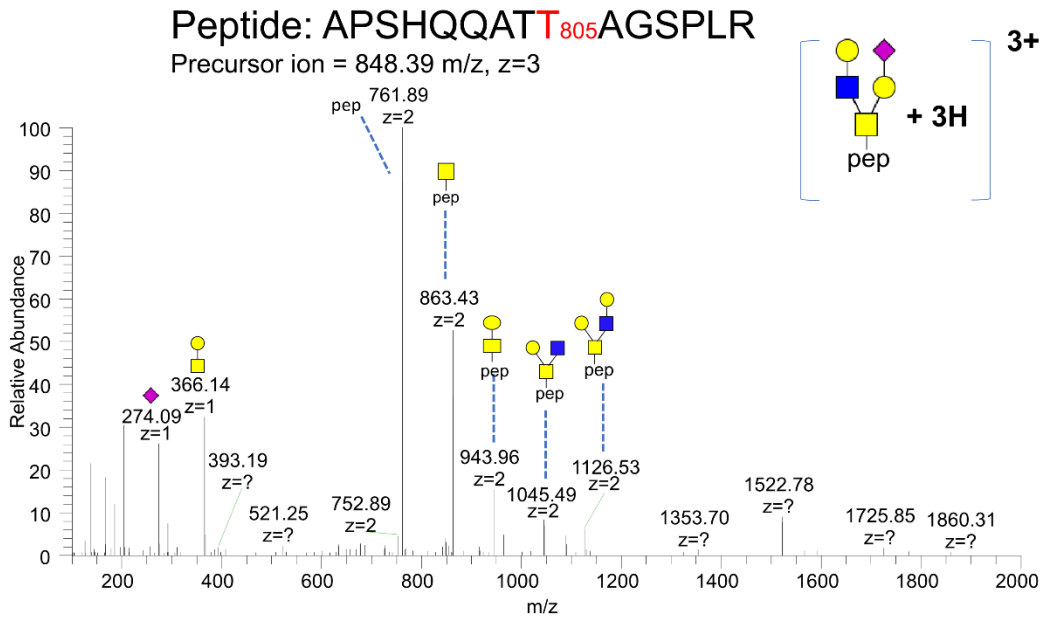


Figure 2.12 Tandem MS annotation of O-glycopeptide APSHQQATT805AGSPLR.

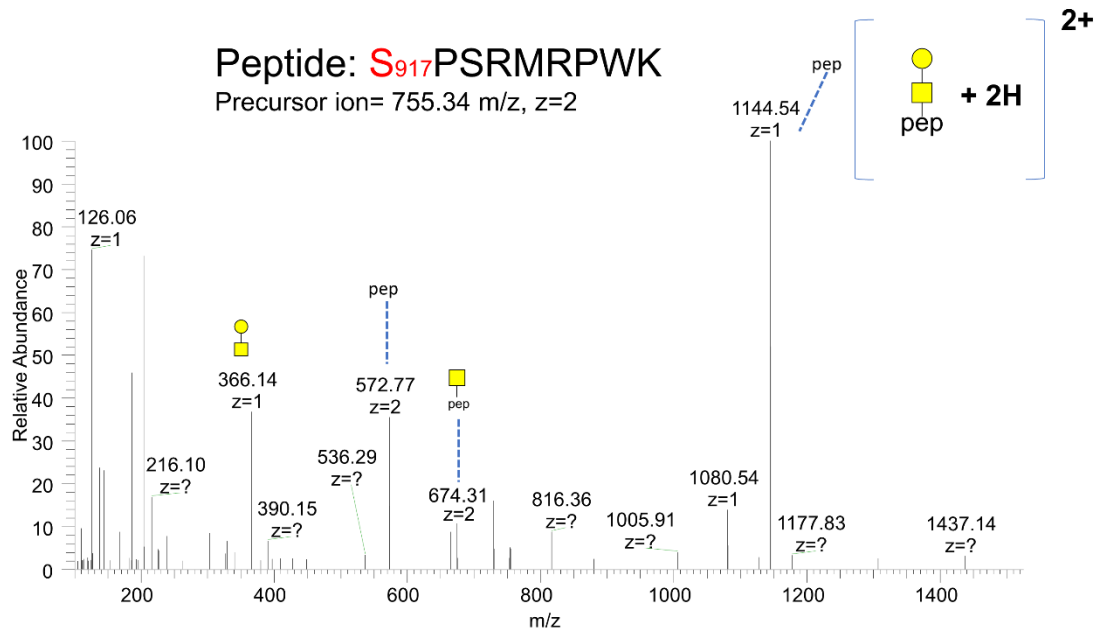


Figure 2.13 Tandem MS annotation of O-glycopeptide S917PSRMRPWK.

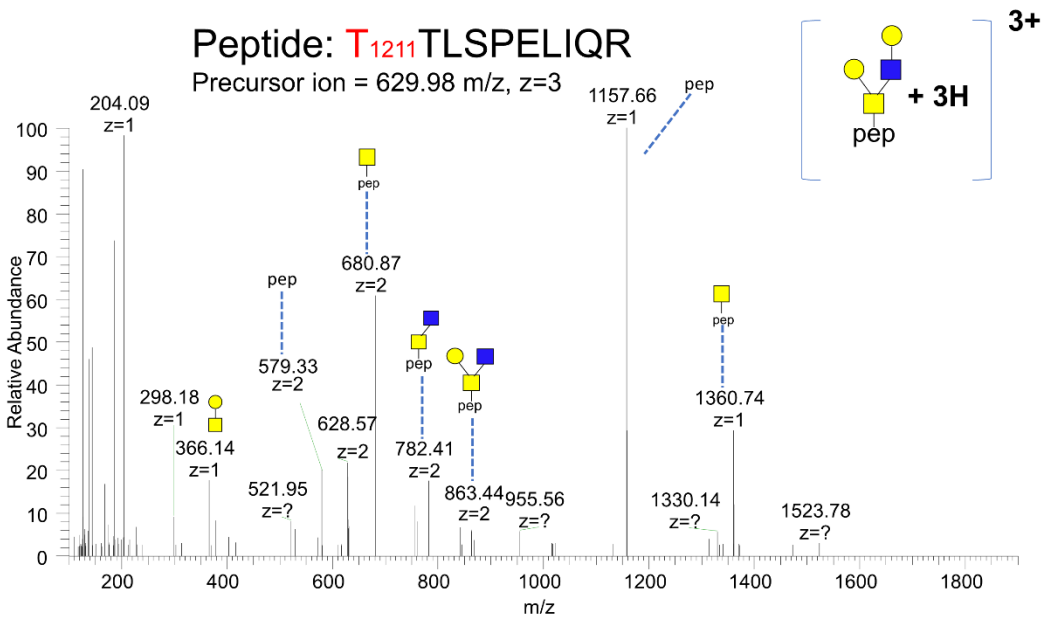


Figure 2.14 Tandem MS annotation of O-glycopeptide T1211TLSPELIQR.

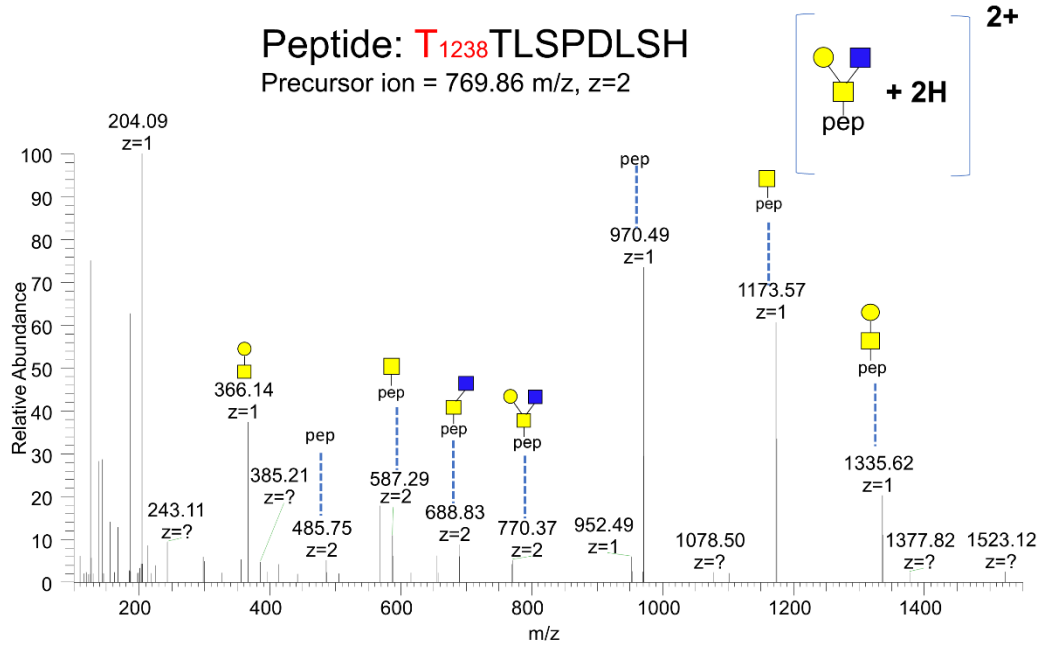


Figure 2.15 Tandem MS annotation of O-glycopeptide T1238TLSPDLSH..

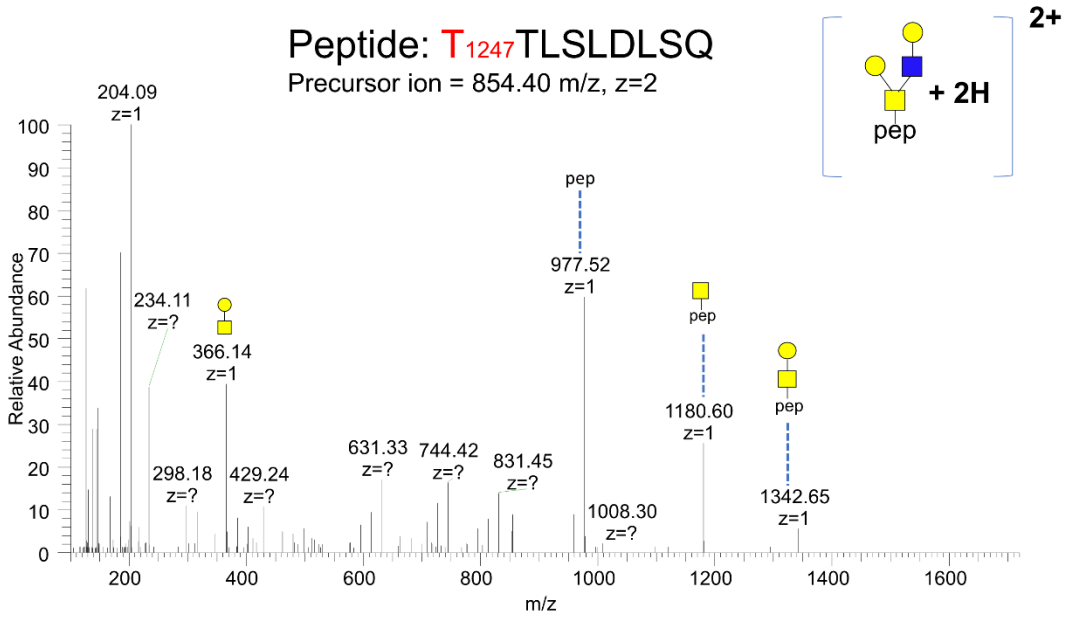


Figure 2.16 Tandem MS annotation of O-glycopeptide T1247TLSLDLSQ.

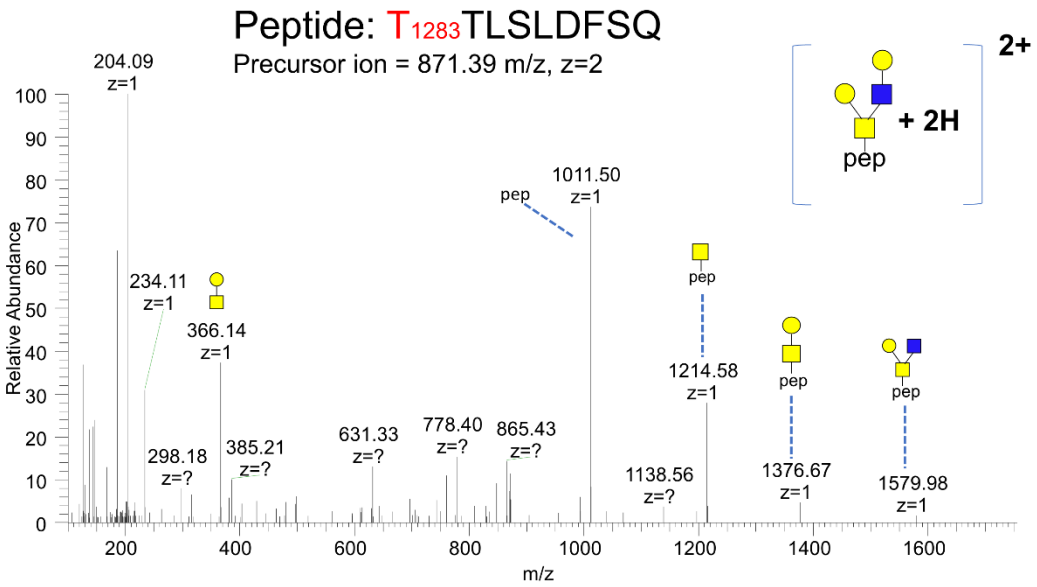


Figure 2.17 Tandem MS annotation of O-glycopeptide T1247TLSLDFSQ.

3 CHAPTER 3. INTEGRATED PROTEOMICS AND GLYCOPROTEOMICS ANALYSES OF COLON CANCER

3.1 Abstract

Colorectal cancer (CRC) is the third most common cancer and the fourth fatal cancer. As current main diagnostic method, endoscopic examination is invasive. Present biomarkers such as CEA and CA-19-9 were lack of sensitivity and specificity for early stage CRC. The shortage of diagnostic methods cause the painful treatments of the CRC patients in the late stages. In recent decades, there is increasing focus on glyco-biomarkers discovery for different diseases through Mass Spectrometry technique. In this study, an integrated glycoproteomics approach was applied to reveal the pathological changes of some CRC-related glycoproteins. two N-glycans have been distinguished with significant differences in 7 distinctly identified N-glycan in this study. On the other hand, among 20 distinctly glycosylated proteins, two N-glycoproteins were discovered with significant difference in terms of glycosylation. The integrated glycoproteomics study provides a comprehensive understanding of glycosylation pattern in colon cancer serum and identified four potential biomarkers for colon cancer.

3.2 Introduction

Colorectal cancer (CRC) is the third most common cancer and the fourth main cause of cancer-related deaths in the world.[1] According to statistics study from National Center for Health Statistics, there is nearly 14790 new cases of CRC and 53200 related deaths in United States during 2020.[2] Mortality rates of CRC are extremely high, partly because it is hardly detected during its early stages. The metastatic CRC is gradually developed from benign polyps.[3] As the main diagnosis method of CRC, Endoscopic examination is uncomfortable and associated with risk of mortality. Several biomarkers have been identified for CRC, such as the carcinoembryonic antigen (CEA)

and the carbohydrate antigen CA-19-9, both of which are carbohydrate or its conjugate.[4] However, these biomarkers has low specificity to CRC and cannot be used for early diagnosis. Carcinoembryonic antigen is generally used for auxiliary diagnosis and has low sensitivity (30%-40%) for early stage CRC,[5] while CA19-9 only shows 8%-17% sensitivity for stage I and II CRC patients.[6] Therefore, the demand of early diagnostic method is urgently.

As a form of post-translational modification (PTM), glycosylation is an enzymatic process that attaches carbohydrate to proteins or other organic molecules.[7] Importance of glycosylation in cancer research has been gained through recent decades. Tumor specific antibodies were directed against glycan epitopes, mostly presented on glycoproteins.[8] Increased evidences prove that glycosylation potentially impact on CRC. For example, α2,6-sialyltransferase I (ST6Gal-I), an enzyme which can produce α2,6-sialylated N-acetyl lactosamine (Sia6LacNAc), was reported to be higher expressed in CRC.[9] Another oligosaccharide called Sialyl Lewis x (sLex) is a selectin ligand, which is overexpressed in CRC patients with poor survival.[10] GnT-V, is an glycotansferase that synthesize β-1,6-GlcNAc on N-glycan. According to Guo et al reported, depletion of GnT-V can reduce the compartment of CRC stem cells and delay the cancer progression in mouse model.[11]

Omics techniques have been applied to clinical cancer research for years, new therapeutic targets of different cancers were discovered by Omics in recent years.[12, 13] In a large scale proteomics study of early-stage hepatocellular carcinoma (HCC), sterol O-acyltransferase 1 (SOAT1) was found to high express in HCC.[13] Treatment of avasimible to inhibit the SOAT1 could notably reduce the size of tumor of HCC in mouse model.[13] In a multi-omics study of colon cancer, proteomics, genomics and transcriptomics were combined to reveal some new therapeutic targets of CRC.[12] Increased glycolysis in tumor was found to be associated with decreased CD8 T cell.[12] As a new developed technique, glycoproteomics has also been widely applied in cancer

research. In an integrated glycoproteomics study of small cell lung cancer (SCLC), fucosylated protein Paraoxonase 1 (PON1) was found and validated as a serum biomarker of SCLC.[14] Otherwise, Hui Zhang and coworkers found increased fucosylation and fucosyltransferase in prostate cancer cells by proteomics and glycoproteomics.[15] There are certain glycoproteomics studies of gastrointestinal-related cancers. For example, a lectin enrichment based N-glycoprotein analysis of HT-29 human colon cancer cells was performed to discover the N-glycan changes in CRC.[4] Celso A. Reis et al enriched the O-glycoproteins by lectin and analyze them for gastric cancer cell lines.[16] However, as increasing attention is attracted to the glycosylation of cancer, a systematic study of CRC glycosylation on clinical samples is still lack.

Therefore, we present the integrated glycoproteomics study of CRC on colon cancer serum. Glycopeptides was enriched from patients' serum and fully studied by mass spectrometry techniques. Occupancy of different glycans on high abundant glycoproteins were fully observed. Quantitative glycoproteomic approach was utilized to study glycosylation changes between CRC patients and healthy human serum. The altered glycosylation patterns among glycoproteins in CRC serum have potential impact on biomarker and therapeutic targets discovery.

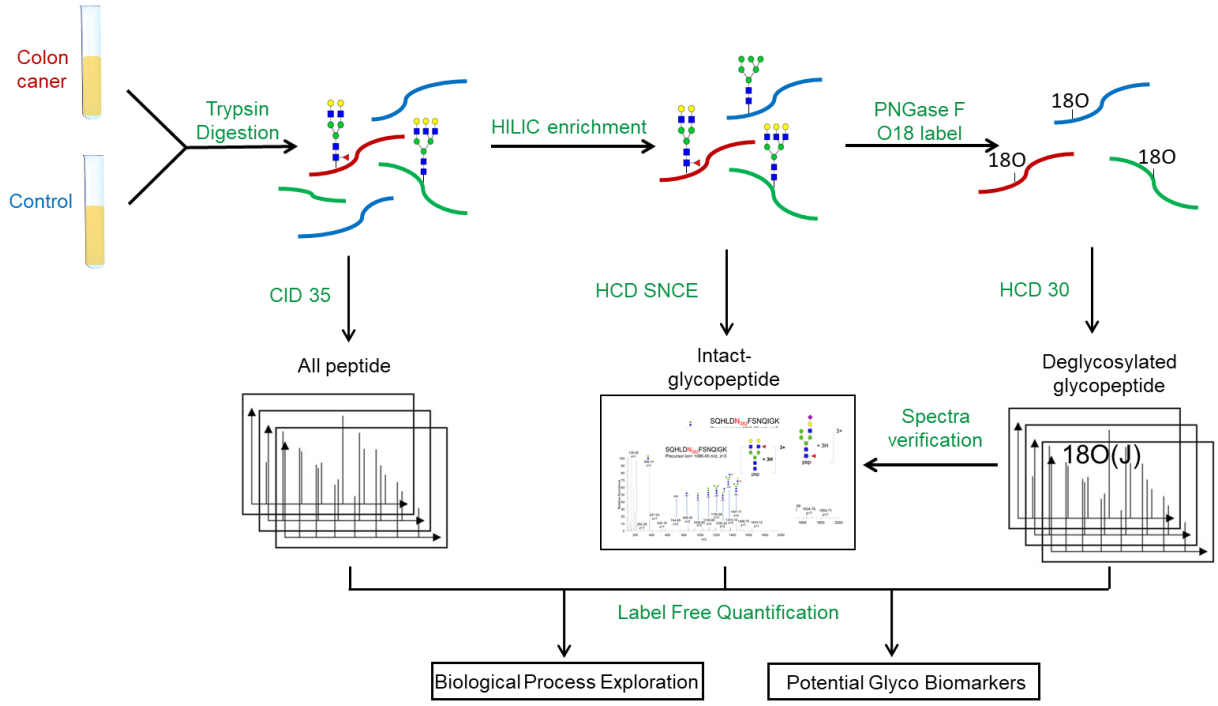


Figure 3.1 Strategy for comprehensively glycoproteomics analysis of colon cancer patients' serum.

3.3 Experimental

3.3.1 Chemical and materials

1,4-Dithiothreitol (DTT), Iodoacetamide (IAA) and trifluoroacetic acid (TFA), Ammonium bicarbonate (NH4HCO3, ABC), Urea , Water-^{18}O with 97 atom % ^{18}O, standard proteins Fetuin from fetal bovine serum, the healthy human serum (from platelet poor human plasma)was purchased from Millipore Sigma (St. Louis, MO). HyperSep C18 SPE Cartridge, Pierce BCA protein assay kit and Nunc MicroWell 96-Well Optical-Bottom Plates was acquired from Thermo Fisher Scientific, Inc. (Waltham, MA). iSPE® HILIC SPE Cartridge was purchased from The Nest Group Southborough, MA). Sequencing grade porcine trypsin was purchased from Promega (Madison, WI). Peptide-N-glycosidase F (PNGase F) was bought from New England Biolabs (Ipswich, MA, USA). Deionized water (DI water) was produced by a Milli-Q A10 system from

Millipore (Bedford, MA). UltraPure 1M Tris-HCl pH 8.0, Optimal LC-MS grade quality acetonitrile (ACN), formic acid (FA) and water were purchased from Thermo Fisher Scientific, Inc. (Waltham, MA).

Serum samples from with 10 colon cancer patients and 10 healthy people were provided by Georgia Cancer Center at Augusta University and stored at -80 °C until use. The protocol for serum sample preparation was approved by the Ethics Committee of Augusta University and was performed in accordance with the Helsinki Declaration. All participants gave written informed consents.

3.3.2 Serum protein concentration measurement

Protein concentration of all samples were measured by BCA assay. BCA working reagent (WR) was prepared by mixing 50 parts of BCA Reagent A with 1 part of BCA Reagent B. Serum samples were diluted by DI water with 50 dilution factor. BSA standards with 0.125, 0.25, 0.5, 0.75, 1, 1.5, 2 mg/ml concentration were used to acquire the standard curve of BCA assay. 25 μL of each diluted sample and 200 μL of WR were added to 96 well plate and mixed. The plate was covered and incubated at 37 °C for 30 minutes. The plate was cooled and measured at 560 nm by Victor UV/ Vis/FluorescencePlate Reader.

3.3.3 Protein digestion

The protein digestion of serum was performed by in-solution digestion process.[17] 20 μg of Bovine Fetuin was added to each sample as internal control. Human serum contains 1mg protein was mixed with UA buffer (8M urea, 50 mM Tris-HCl, pH 8.5). After adding 5 mM DTT, the sample was denatured at 95 °C for 10 min. Then 20 mM IAA was added, the solution was incubated for 30 min in the darkness at room temperature. After that, another 20 mM DTT followed by 1 mL of 50 mM ABC were added to the solution. At last, 20 μg of trypsin was added into the

solution and incubated at 37 °C for 15 hours. 1% Formic Acid was used to denature the trypsin. The digested protein was desalted by previous described.[18] Briefly, C18 SPE cartridge was activated by ACN and equilibrated by 0.1 % TFA in water. The sample was loaded to equilibrated column for three times, then the column was washed with 12 mL 0.1 % TFA in water and eluted by 8 mL of 80% ACN with 0.1 %TFA. The eluted solution was dried by Rotovap at 30 °C and lyophilized for next step.

3.3.4 Glycopeptide enrichment

The glycopeptides enrichment was performed using iSPE HILIC cartridge with the procedures of ZIC-HILIC. Briefly, the HILIC cartridge was pre-washed with 600 µL of 0.1% FA and equilibrated with 2 mL of 80% ACN containing 5% FA. Peptide samples dissolved in 400 µL of 80% ACN containing 5% FA were loaded onto the cartridge for three times. Then the column was washed with 2.5 mL of 80% ACN containing 5% FA. The glycopeptides were eluted using 3 mL of 0.1% FA. The eluted solution was lyophilized and stored at -20 °C for next step.

3.3.5 N-Glycosylation site label by ^{18}O

Glycopeptides enriched from 200 µg protein of each human serum sample were used for N-glycosylation site labelling. N-glycosylation sites were labelled with ^{18}O by using previous reported method[19]. Glycopeptides were dissolved in 10µL of 25 mM ABC in water-^{18}O, then 500 units of PNGase F was added and mixed with the solution. The solution was sealed and incubated at 37 °C for 4 hours. The endoglycosidase enzyme was denatured at 95°C for 5 min, the labelled glycopeptides were lyophilized and stored at -20 °C for next step.

3.3.6 Nano LC-MS/MS analysis

Nano RP LC-MS experiments were performed on an LTQ-Orbitrap Elite mass spectrometer (Thermo Fisher) equipped with EASY-spray source and nano-LC UltiMate 3000

high-performance liquid chromatography system (Thermo Fisher). The separation was performed on an EASY-Spray PepMap C18 Column (75 μm × 50 cm, 2 μm, ThermoFisher, US) with a linear gradient from 3% to 40% mobile phase B for 120 mins at a flow rate of 300 nL/min (mobile phase A: 2% ACN, 98% H2O, 0.1% FA; mobile phase B: 80% ACN, 20% H2O, 0.1% FA). The mass spectrometer was operated in the data-dependent mode. A full-scan survey MS experiment (m/z range from 375 to 2000; automatic gain control target, 1,000,000 ions; resolution at 400 m/z, resolution at 60,000; maximum ion accumulation time, 50 ms) was acquired by the Orbitrap mass spectrometer. The AGC target for MS1 was set for 1000000 for MS1 in 50 ms maximum injection time (IT). The AGC target for MS/MS was 50000 (resolution of 15000). The ten most intense ions were selected and fragmented by higher-energy C-trap dissociation (HCD 30). An optimized Stepped Normalized Collision Energy (SNCE) of 30 ± 15% NCE was applied to intact-glycopeptide fragmentation.

3.3.7 Data analysis

The pFind 3.0 software was used to identify proteins and N-glycosylation sites. Uniprot Swiss-Prot human fasta file (2020_03 release, 20365 reviewed entries) was used for database search.[20] In amino acid sequence of the fasta file, N in sequence N-X-S/T/C (X≠P) was replaced with J by using the python program as we previous described.[21] The raw file was identified using pFind software to perform the database search against the modified fasta file. In pFind software, J was defined as asparagine with the same monoisotopic mass, and the variable modification “18O” was set on J as N-glycosylation site label. Other Searching parameters were set as common proteomics. Fixed modification: carbamidomethyl (Cys); Variable modifications: 18O (J), deamination (N, Q) and oxidation (M). Trypsin was chosen as the enzyme, and two missed cleavages were allowed. The mass tolerance for the precursor ions and the fragment ions was set

to 20 ppm and 0.5 Da. False discovery rate (FDR) based on decoy of 1% was applied to all data sets at the peptide level as previous reported.[22] Then pBuild was used to remove protein entries without modification 18O (J). The target peptides, containing modification 18O (J), were used to build up an experimental spectral library (ESL) for intact-glycopeptide search.

The GPQuest software was used to identify intact-glycopeptides.[23] In the glycan database, human N-glycan sequences following the motif N#H#F#A#G# where "N" is for N-Acetylglucosamine, "H" for Galactose, "F" for fucose, "A" for Neu5Ac and "G" for Neu5Gc. For intact-glycopeptides identification, the raw data was firstly converted to mzML file using MSConvert.[24] The oxonium ions were used to differentiate the glycopeptide and peptide. The MS/MS spectras containing two or more of oxonium ions, including 126 $[HexNAc-C2H6O3]^+$, 138$[HexNAc-CH6O3]^+$, 145 $[Hex- H2O]^+$, 168 $[HexNAc-2H2O]^+$, 186 $[HexNAc- H2O]^+$, 204 $[HexNAc]^+$, , 274 $[Neu5Ac-H2O]^+$, 292 $[Neu5Ac]^+$ or 366 $[HexHexNAc]^+$ were considered as glycopeptide spectras.[25] Each oxonium-ion-containing MS/MS spectra was compared with the corresponding experimental spectral library (ESL). The mass tolerance for the precursor ions and the fragment ions was set to 10 ppm and 50ppm. The shift from precursor mass to monoisotopic mass of theoretical peptide was calculated to annotate the glycan composition, then exact match with mass tolerance of 10ppm was searched against the human glycan database. The results were further filtered by applying 1 % FDR. The results of glycosylation-site and intact-glycopeptide were further processed by python program for data analysis.

3.4 Results and discussion

3.4.1 Sample information and protein concentration

The entire study included 20 subjects including 10 patients with colon cancer and 10 controls with same race and gender. The average age at time of sample collection was 67.2 for

patients and 67.3 for controls. The patients' samples contain seven cases of stage III, one case of stage II, two cases of stage IV.

Table 3.1 Sample information of human serum.

	Colon cancer	Control
Gender	5/5 male/female	5/5 male/female
Age	67.2±8.6	67.3±7.86
Stage	1/7/2 II/III/IV	None

The protein concentrations of the serum samples were acquired through BCA assay. 2 μL of each serum was diluted by DI water with 50 dilution factor. Each diluted sample was incubated with WR buffer in 3 individual wells (25 μL /each). After measurement, the average of three repeated reactions were calculated as absorbance results. Bicinchoninic acid (BCA) assay is a detergent-compatible formulation that depends on the colorimetric reaction between bicinchoninic acid and cuprous cation (Cu+). The method has high sensitivity and selectivity on protein concentration measurement.[26] The absorbance at 562 nm shows a linear correlation with the concentration range of the standard protein (BSA). The total protein concentration of soluble proteins derived from colon cancer serum samples are ranged from 54 to 86 mg/mL. Each serum sample containing 1mg of protein was used for subsequent proteomics and glycoproteomics analysis. Standard protein (BSA) with gradient concentrations were applied in BCA assay to acquire the standard curve. (Figure 3.2) For human serum sample, dilution factor was used to calculate the total protein concentration. (Table 3.4)

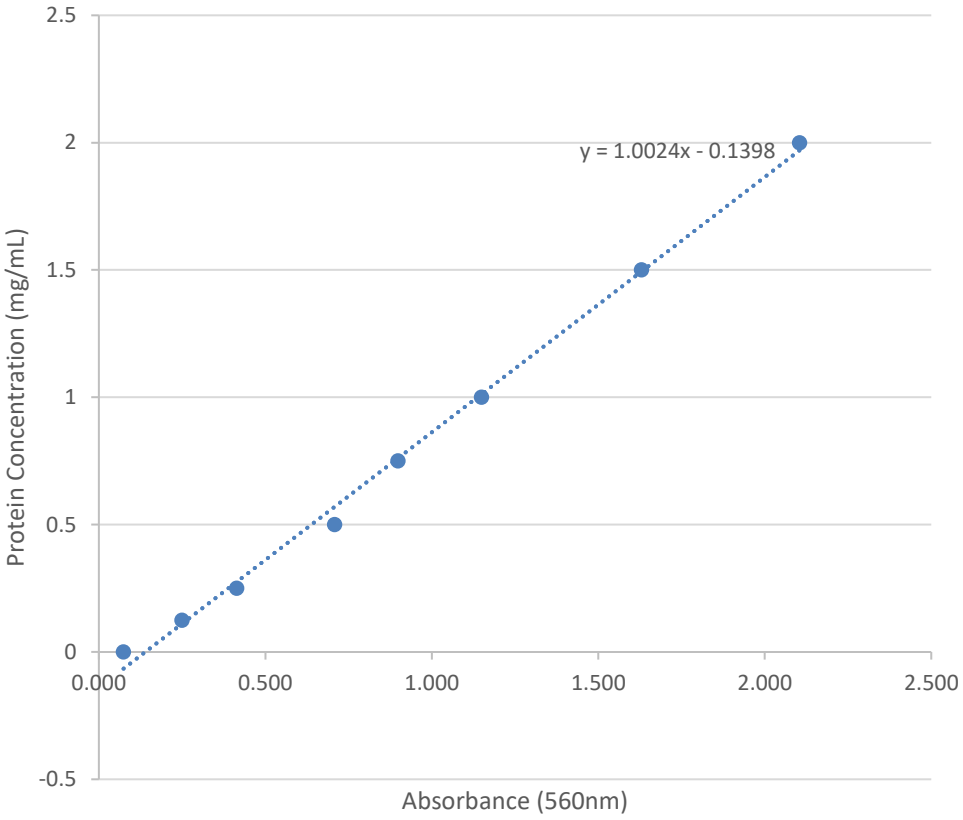


Figure 3.2 Standard Curve of BCA assay for human serum samples' protein concentration calculation.

3.4.2 Identification of intact-glycopeptide by optimized SNCE method

Intact-glycopeptide analysis could take advantage of the fragmentation of glycopeptide under HCD tandem MS method, since its high accuracy and it can produce characteristic tandem MS peaks (oxanium ion) for glycopeptide.[23, 25] It was reported that HCD with stepped normalized collision energy (SNCE) could produce abundant fragment ions for both the glycan and peptide of glycopeptide in a single spectra.[27] In the previous reported method, SNCE-HCD-MS/MS under 20–30–40% could generate most information of glycopeptide in Orbitrap Fusion Tribrid (Thermo Scientific) system.[27] However, the optimized SNCE method was not reported for glycopeptide fragmentation of Orbitrap Elite (Thermo Scientific) system in our lab.

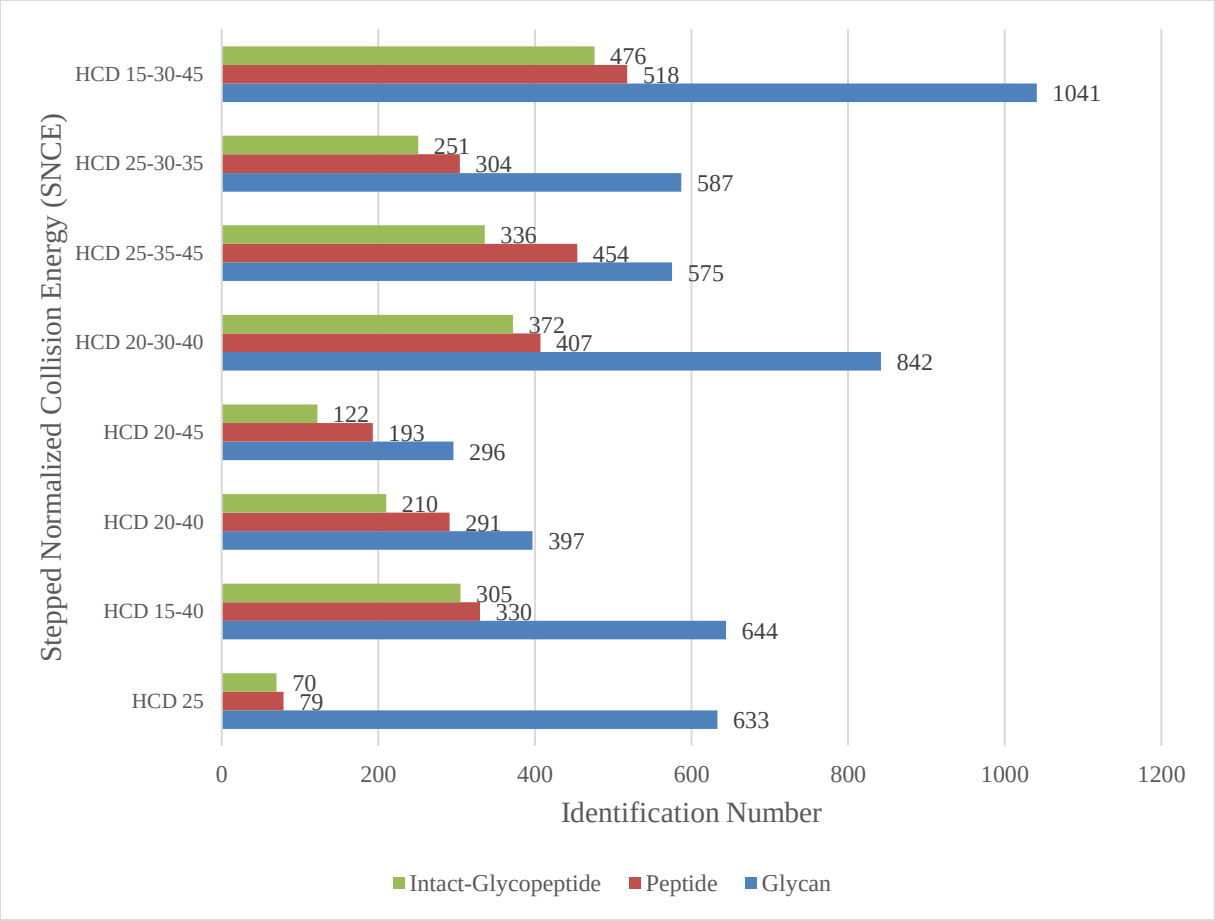


Figure 3.3 Optimization of glycopeptide identification under different normalizaed collision energy methods.

To optimize the SNCE method for our mass spectrometry system, glycopeptide enriched from healthy human serum (Millipore Sigma) was used for optimization. Several NCE, two-step SNCE and three-step SNCE methods were applied in the glycopeptide identification acquire the ideal method. The identification of N-glycopeptides was performed using GPQuest software. As shown in Figure 3.3, most amount of glycopeptides from the same sample can be identified under SNCE of "HCD 15-30-45". The optimized SNCE condition was then applied to the following glycoproteomics study of colon cancer serum.

3.4.3 Colon cancer unique protein and glycosylation site

In this study, totally 381 proteins, 96 glycoprotein with 176 N-glycosylation sites were identified in 3 replicates. Gene Ontology (GO) analysis was processed by using PANTHER classification system.[28] In the PANTHER system, the identified glycoproteins were classified into different group based on their function in biological system. As shown in Figure 3.4 B, 55.1 % of the glycoproteins were immunity protein, which consists of different immunoglobulins. Other major glycoproteins included 13.3 % of protein binding activity modulators, 10.2 % protein modifying enzyme, 5.3 % carrier proteins, 4.4 % cytoskeletal protein, 3.1 % intercellular signal molecule, and other proteins. Through the Gene Ontology (GO) analysis, the proteins mainly interacted and modulated other proteins' activity and played an important role in immune response.

Among the 381 identified proteins, 328 proteins were detected in both colon cancer groups and healthy human groups for 3 repeats. (Figure 3.4 A) The other 51 proteins were refined by following regulation. The proteins identified in 3 replicates of colon cancer serum and none of healthy human serum were defined as colon cancer unique glycoproteins. The proteins identified in 3 replicates of healthy human serum and none of colon cancer serum were defined as healthy

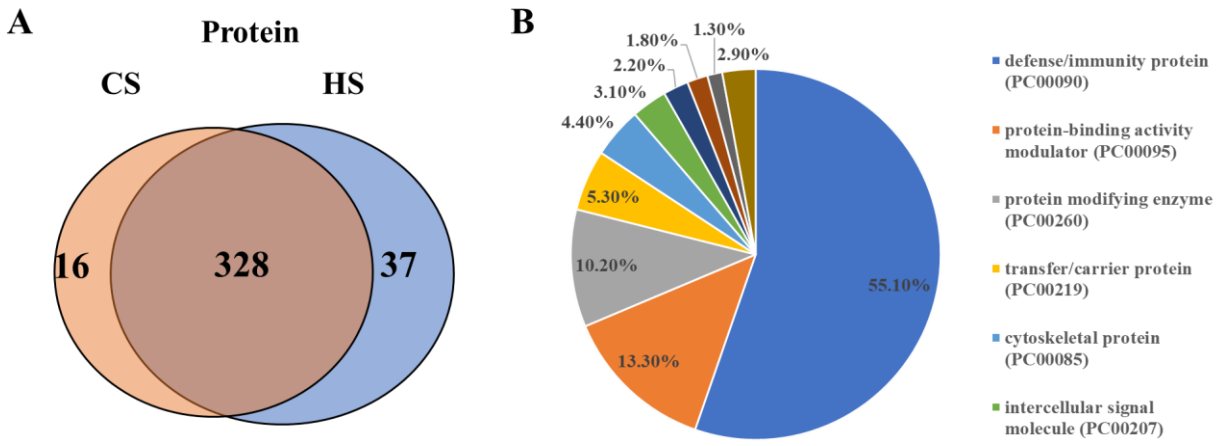


Figure 3.4 Protein function classification of glycoprotein identified from human serum samples.

human unique proteins. Through this regulation, one protein were defined as colon cancer unique proteins, while 4 proteins were define as healthy human unique proteins. Immunoglobulin lambda variable 3-16 was the colon cancer unique protein, which is part of the variable domain of immunoglobulin light chains that participates in the antigen recognition. Three of the 4 healthy human unique proteins were homeobox proteins, and the other one was carbonic anhydrase 1. (Table 3.5) Homeobox were transcription factor, and aberrant expression of homeobox genes in cancer was reported.[29] Moreover, carbonic anhydrase 1 was reported to significantly elevated in non-small cell lung cancer.[30] The homeobox proteins and carbonic anhydrase 1 were only detected in healthy human might indicate the lower expression of these proteins in colon cancer serum.

Among the 96 identified glycoproteins, 71 glycoproteins were detected in both colon cancer groups and healthy human groups for 3 repeats. (Figure 3.5 A) The other 25 glycoproteins were refined by following regulation. The protein should be identified in both groups, the glycosylation sites identified in 3 replicates of colon cancer serum and none of healthy human serum were defined as colon cancer unique glycosylation site. The glycosylation sites identified in 3 replicates of healthy human serum and none of colon cancer serum were defined as healthy human unique glycosylation sites. Through this regulation, two glycosylation sites were define as healthy human unique glycoproteins, including N73 of carboxypeptidase B2 and N308 of plasminogen. (Table 3.5) Carboxypeptidase B2 (CPB2) can inhibit lysis of the fibrin clot and reduce plasminogen activation, previous study showed that pro-CPB2 protein might be associated with poor prognosis of colon cancer.[31] Glycosylation was removed during the activation of pro-CPB2 to CPB2,[32] which correspond to our observation of non-glycosylated CPB2 in colon cancer.. Plasminogen is an inactive serine protease, which can breakdown extracellular matrix during

cancer invasion process.[33] The activation of plasminogen was regulated by some inhibitors and activators, including urokinase plasminogen activator (uPA) and plasminogen activator inhibitors 1 (PAI-1), which were reported to be associated with colon cancer.[34] The absence of plasminogen's glycosylation in our results might be involved in its activation, thereby affect the cancer invasion process. Th unique glycosylation observed on CPB2 and plasminogen in healthy human serum might refer to the decreased of glycosylation in colon cancer, which is of interest to further study.

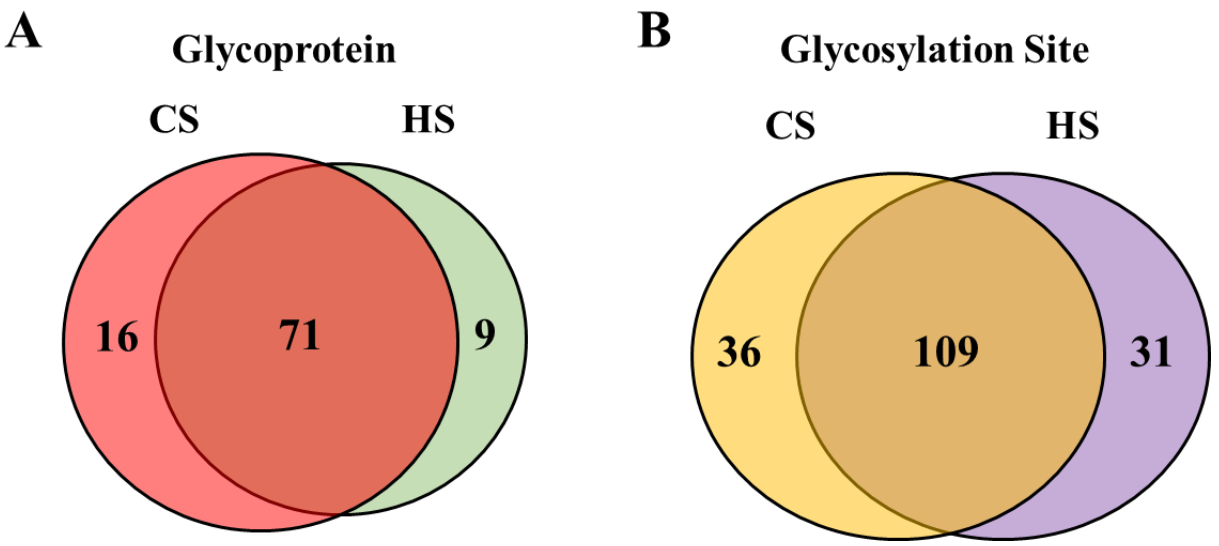


Figure 3.5 Overlap of glycoproteins and N-glycosylation sites identified from colon cancer serum (CS) and healthy human serum (NS).

Besides glycoproteins and N-glycosylation sites, the overall change of glycoform was also observed. In this study, 58 different N-glycoforms were identified in more than 3 samples totally, while 54 N-glycoforms of them were detected in both colon cancer serum and healthy human serum. Among the 54 glycoforms, di-antennary complex type glycan consisted 31 %, tri-antennary complex type glycan consisted 27.6 %, tri-antennary complex type glycan consisted 8.6 %, hybrid type glycan consisted 19 % and high mannose type glycan consisted 13.8 %. High reliable unique glycoform was not found in colon cancer serum or healthy human serum.

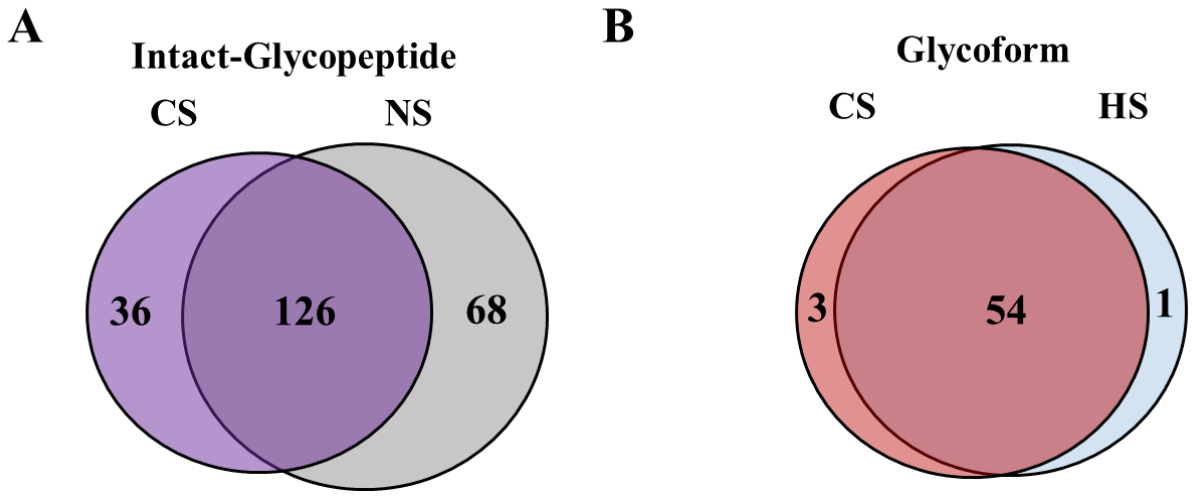


Figure 3.6 Overlap of Intact-glycopeptide and N-glycoforms identified from colon cancer serum (CS) and healthy human serum (NS).

3.4.4 Glycan type percentage comparison

N-glycans are usually classified into high-mannose type, hybrid type and complex type. Thus the percentage comparison between different glyan type on these glycosylation site was performed. The fucosylation and sialylation were also compared. Fucosylation and sialylation were known to be associated with cancer. Fucosylation is important in blood group, immune system and signaling pathways,[35] and the fucosyltransferase 3 and/or 6 was proven to be highly related with metastasis of colon cancer.[36] Moreover, sialylation could affect the function of TNF-related apoptosis-inducing ligand (TRAIL), which could kill tumor cells.[37]

Firstly, the overall N-glycan type percentage between colon cancer serum and healthy human serum was compared by spectra counting method. The spectra number of each glycan type was divided by total spectra count number of all N-glycan. As shown in Figure 3.7, N-glycans in human serum contained 15~20 % high-mannose type, 28 % hybrid type, 40 % di-antennary complex type and 11~15% tri-antennary complex type. Nearly 50 % glycans were fucosylated, and 15 % glycans were sialylated. The percentage of hybrid type , di-antennary type , fucosylation

and sialylation did not change. The tri-antennary complex type decreased 3 %, and the high-mannose type glycan increased 6 % in colon cancer.

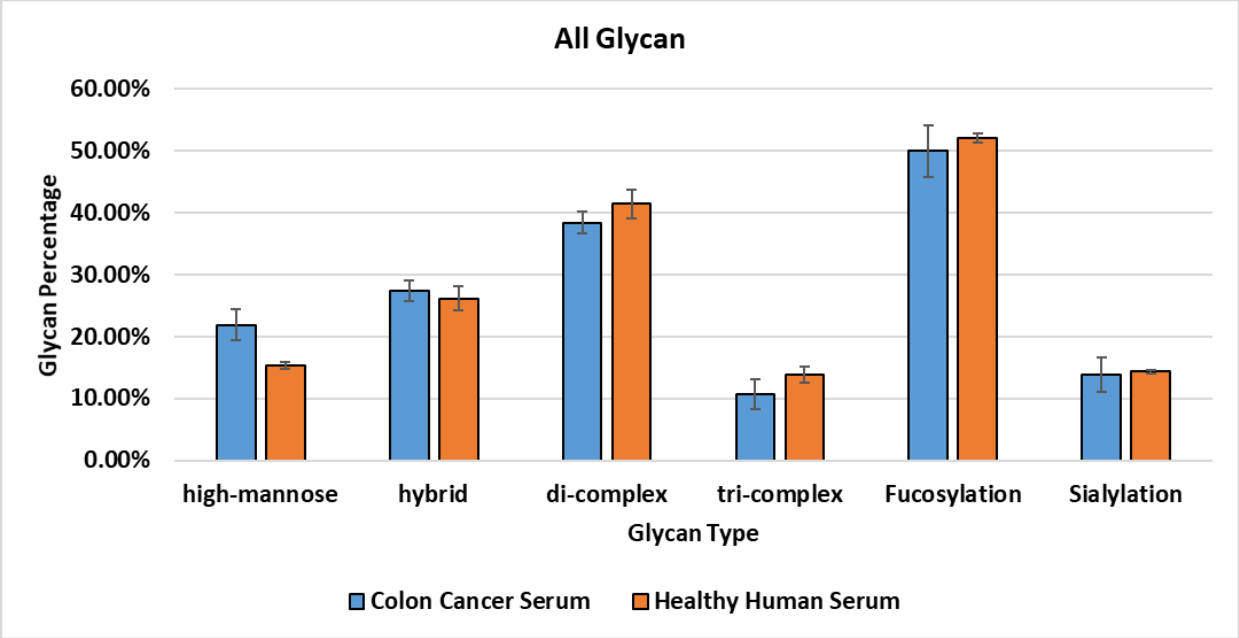


Figure 3.7 Glycan type percentage of all N-glycan in the human serum.

For glycoprotein with high abundant glycopeptide, different glycoforms could be detected in one glycosylation site. The spectra number of each glycan type on the same site was divided by total spectra count number of the same peptide. As shown in Figure 3.8, percentage of each glycan type were calculated and compared on totally 7 glycosylation sites, including 5 sites from immunoglobulin proteins. The overall distribution of different glycoforms on each glycosylation site was discussed. As shown in Figure 3.8, for Immunoglobulin G1 (IgG1) and Immunoglobulin G2 (IgG2), all of their glycosylation were on the similar position of the constant heavy 2 (CH2) domain. From our results, all of these IgG had 36~42 % hybrid type glycan, around 60% complex type glycan. Among the 60% complex type, IgG1 and IgG2 had 45~51 % di-antennary glycan and 10~15 % tri-antennary glycan. Moreover, IgG1 had around 62 % fucosylation, 0~4 % sialylation, and IgG2 had 70~80 % fucosylation, 2~3 % sialylation. The Immunoglobulin M (IgM) showed

similar trending with Immunoglobulin G in terms of glycan types, but its sialylation (12~13 %) was considerably higher than IgG. The Immunoglobulin A (IgA) had two high abundant glycosylation site (N131 and N205). N131 of IgA showed high heterogeneity, which contained around 9% of high-mannose type glycan, 18 % of hybrid type glycan, 34 % bi-antennary complex type glycan and 38 % tri-antennary complex type glycan. Moreover, it had 33 % fucosylated glycan and 6 % sialylated glycan. Differently, high-mannose type glycan was not detected on N205 of IgA, and fucosylation (95 %) and sialylation (18 %) were significantly higher than N131. Two glycoproteins other than Immunoglobulin were Complement C3 (CO3) and haptoglobin. Haptoglobin showed high heterogeneity, it contained around 12 % of high-mannose type glycan, 59 % of hybrid type glycan and 25 % bi-antennary complex type glycan. Haptoglobin had low fucosylation (26 %) and sialylation (12 %). Differently from haptoglobin, the glycosylation site N85 of CO3 showed low heterogeneity, it included majority of high-mannose type (89 %) glycan and hybrid type (11%) glycan, it also had around 11% sialylation.

Differences were observed among these glycosylation sites. In colon cancer serum, fucosylation of IgA2 increased from around 89% to 100 %. Differences of sialylation were also inspected among these glycosylation sites. In colon cancer serum, sialylation of N131 and N205 were slightly decreased around 3%, while sialylation of all other glycosylation sites were lightly increased 1~4 %. Besides, significant increase of hybrid type and decrease of tri-antennary complex type glycan was found in IgM. Differently, decreased hybrid type and increased di-antennary complex type glycan was observed in haptoglobin.

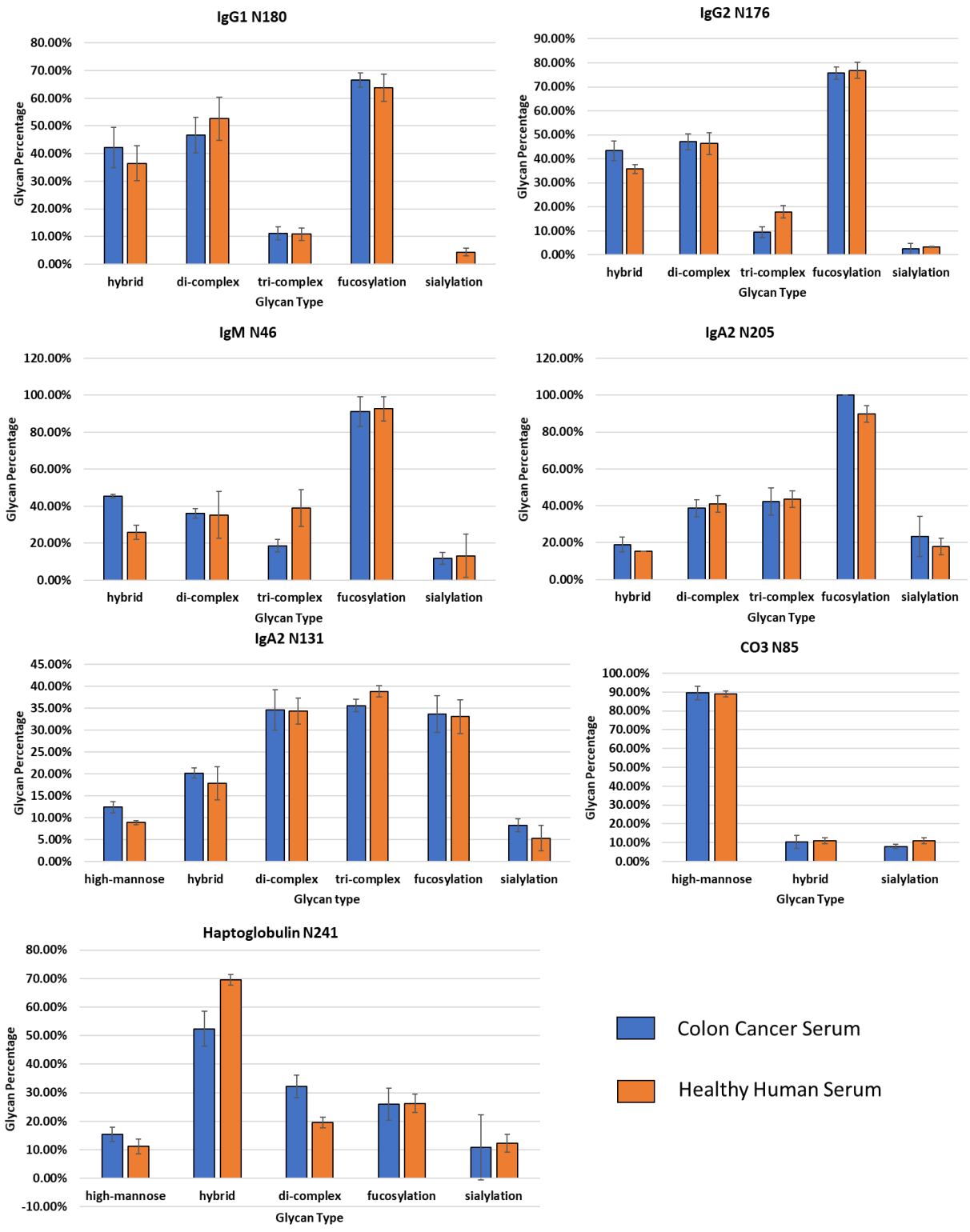


Figure 3.8 Percentage of different types glycan on specific glycosylation sites.

Table 3.2 Significant changed glycan type percentage in all glycan and three glycoproteins.

Glycosylation Site	Glycan Type	Colon Cancer	Healthy Human	Change	Trend
All Glycan	High-mannose	21.83%	15.32%	6.51%	↑
IgM (N46)	Hybrid	45.35%	25.76%	19.59%	↑
	Tri-antennary Complex	18.60%	38.96%	-20.36%	↓
IgA2 (N205)	Fucosylation	100.00%	89.74%	10.26%	↑
Haptoglobin (N241)	Hybrid	52.38%	69.51%	-17.13%	↓
	Di-antennary complex	32.25%	19.51%	12.74%	↑

In overall N-glycan, high-mannose type glycan was reported to be elevated in tumor cells compared with normal cells,[38] which is identical to our results. IgM is the first line of defense in the immune system, it mediates B cell development, promotes the removal of apoptotic cells and regulates the elimination of cancer cells.[39] In ovarian cancer serum, the high-mannose type glycan on N439 of IgM was differentially expressed.[40] While this change was not observed in our study, the 20 % difference of hybrid type and tri-antennary complex type glycan between colon cancer and healthy human provided glycosylation alteration information on N46 of IgM. IgA can recruit more neutrophils (cytotoxic cells) than IgG antibodies, thereby IgA has higher signaling capacity of tumor cells.[41] Most glycosylation study of IgA was focus on IgA1,[42] but the 10 % elevation of fucosylation in N205 of IgA2 raised the importance of its glycosylation in cancer. Haptoglobin is a secreted acute phase protein that is mainly produced in liver. Altered fucosylation in haptoglobin was observed in pancreatic cancer,[43] while 17 % decrease of hybrid type and 12 % increase of di-antennary complex type glycan were found in colon cancer from our results. As these

glycoproteins played important role in cancer, the mechanism of glycosylation change on them were worth further study.

3.4.5 Relative quantitative analysis of protein and glycoprotein

The 328 proteins and 71 glycoproteins identified in both colon cancer serum and healthy human serum were further compared through label-free quantitative analysis. Label-free quantitative analysis method has been widely used in proteomics.[44] There are two different method in label-free quantitative proteomics, peak area measurement and spectral counting.[45] The first method is integrating the peak area of extracted ion chromatograph for target peptides, the second method is counting the mass spectra numbers for target peptides.[44] In our study, spectra count method was used for glycoprotein relative quantitation. Standard glycoprotein Bovine Fetuin was used as internal standard, and same amount of Bovine Fetuin was added to the serum before sample preparation process. The spectra count of each glycoprotein was normalized by spectra count of Bovine Fetuin, then the normalized numbers of colon cancer and healthy human were compared. The quantitative comparison was done by using pooled serum of colon cancer comparing with pooled serum of healthy human with three replicates. This strategy was widely used in biomarker study by proteomics.[46, 47]

Firstly, T test and fold change were applied to find significant change of proteins and glycosylation between colon cancer and healthy human. The proteins or glycoproteins fulfilled the requirement that fold change > 1.5, p value < 0.05 were selected for next step analysis. (Figure 3.9) By applying this strategy, totally 24 proteins were found to have significant difference in glycosylation. (Table 3.6) Among the 24 proteins, 18 proteins had higher abundance in colon cancer, while 6 proteins had lower abundance in colon cancer. Abundance of 9 proteins in colon cancer were 2 to 6 fold higher than healthy human, while abundance of other 10 proteins in colon

cancer were 1.6 to 2 fold higher. Both of 6 protein with negative fold change had larger than 2 fold change value. On the other hand, the glycoproteins from colon cancer serum and healthy human serum quantitatively compared by using the spectra count of ^{18}O labelled glycopeptides. Totally 10 glycoproteins were found to have significant difference in glycosylation. (Table 3.7) Among the 10 glycoproteins, 9 glycoproteins had higher glycosylation in colon cancer, while only one glycoproteins had lower glycosylation in colon cancer. Glycosylation of 7 glycoproteins were 2 to 5 fold higher, while glycosylation of other two glycoproteins in colon cancer were 1.6 to 2 fold higher than healthy human. The glycoprotein with negative fold change had less than 2 fold change.

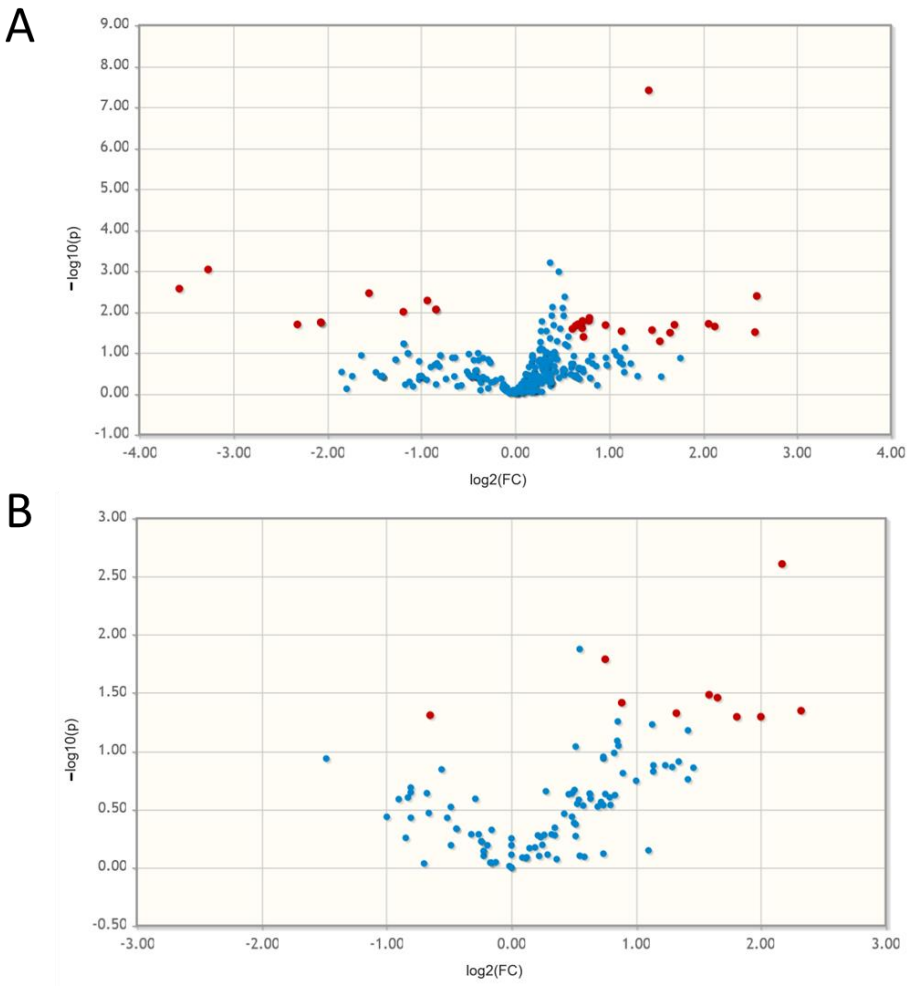


Figure 3.9 Volcano plot of quantitative fold change (colon cancer/ healthy human) of identified proteins (A) and glycoproteins(B). (Fc:Fold change>1.5, p<0.05)

Since glycopeptides were enriched from digested proteins by hilic enrichment method, the quantity of glycopeptide would be changed with the quantity of proteins. Therefore, the protein fold change and glycoprotein fold changed were combined to find the change of glycan occupancy. As listed in Table 3.3, by comparing the change of protein quantity with the change of its glycosylation, the change of glycan occupancy could be concluded.

Table 3.3 Glycoproteins with significant changed glycan occupancy in colon cancer,

Protein Name	**Protein log2Fc**	**Glycosylation log2Fc**	**Glycan Occupancy**	**p value**
Alpha-1-antichymotrypsin	-0.26878	0.75207	↑	0.016352
Plasma kallikrein	-0.43313	0.88452	↑	0.038611
Tetranectin	0.023922	1.3219	↑	0.047421
Apolipoprotein A-II	0.66693	1.585	↑	0.032955
Angiotensinogen	0.71903	1.6521	↑	0.034908
Complement component C9	-0.18247	2	↑	0.050877
Vitamin K-dependent protein S	**-0.52251**	**2.1699**	↑	**0.00249**
Apolipoprotein A-I	-0.32858	2.3219	↑	0.045174
Apolipoprotein M	2.0594	0.80735	↓	0.019752
von Willebrand factor	2.5535	0.66297	↓	0.031502
Cholinesterase	**2.5727**	**-0.58**	↓	**0.004147**

In previous report, statistically significant findings with $p < 0.05$ would result in high false positive rates, the threshold $p < 0.005$ could significantly improve the reproducibility of scientific research.[48] Therefore, this threshold was applied in our study, the p value of Vitamin K-dependent protein S and Cholinesterase were less than 0.005. The potential of them to be biomarkers was further discussed. Cholinesterase (ChE) is an esterase that lyses choline-based esters, it is involved in cellular proliferation and differentiation.[55] The activity of ChE was reported to influence the metastasis of prostate cancer,[56] and ChE genes were amplified in leukemias and ovarian

carcinomas.[57] Glycosylation of acetylcholinesterase (AChE) was proven to decline in human breast cancer through lectin binding method.[58] Although the glycosylation of ChE was widely reported, the importance of its glycosylation in cancer was not brought to the forefront as well as AChE.[59, 60] The significant decrease of ChE glycosylation in colon cancer from our results is of interest to further study.

Vitamin K-dependent protein S (PS) is an anticoagulant plasma protein, but it does not show anticoagulant functions in T cells and tumor cells.[49] In these cells, PS activates receptors of protein–tyrosine kinase (PTK), including Tyro3, Axl, and Mer receptors (TAMR). TAMRs signaling have an important pivotal role in cancer development and progression.[50] Because the overexpression of PS and TAMR were observed together in numerous types of cancers,[51-53] a hypothesis model was proposed that inhibition of TAMR activity through using PS antagonist may suppress tumor progression and metastasis.[54] Binding of PS to TAMRs expressed on the surface of tumor cells and inhibition of the host defense system response may help cells stick together and survive. In our results, the glycan occupancy of PS was significantly increased. As glycosylation plays important role in binding and protein-protein interaction, the elevated glycosylation in PS may enhance its interaction with TAMR, thereby promote the survive of tumor cells.

3.5 Conclusion

In conclusion, we comprehensively identified 328 proteins, 109 N-glycosylation sites and 54 different N-glycans in colon cancer and healthy human serum samples. We identified 5 unique proteins and 2 unique glycopeptides in colon cancer or healthy human, which may indicate their concentration change in colon cancer. In percentage analysis of N-glycan type, increased high-mannose type glycans were observed among all glycans. Percentage change of other glycan types

were detected in IgM, IgA2 and Haptoglobin. In quantitative analysis of glycan occupancy, 3 glycoproteins had lower N-glycan occupancy, and 8 glycoproteins had higher N-glycan occupancy in colon cancer. After further improve the threshold of p value to 0.005, cholinesterase and vitamin K-dependent protein S were found to have significant decreased and increased glycan occupancy individually. The two identified glycoproteins were highly associated with cancer progression, the importance of their glycosylation worth further study.

3.6 Reference

1. Siegel, R. L.; Torre, L. A.; Soerjomataram, I.; Hayes, R. B.; Bray, F.; Weber, T. K.; Jemal, A. J. G., Global patterns and trends in colorectal cancer incidence in young adults. **2019,** *68* (12), 2179-2185.
2. Siegel, R. L.; Miller, K. D.; Jemal, A. J. C. A. C. J. f. C., Cancer statistics, 2020. **2020,** *70* (1), 7-30.
3. Tierney, R.; Ballantyne, G.; Modlin, I. J. S., gynecology; obstetrics, The adenoma to carcinoma sequence. **1990,** *171* (1), 81-94.
4. Vercoutter‐Edouart, A. S.; Slomianny, M. C.; Dekeyzer‐Beseme, O.; Haeuw, J. F.; Michalski, J. C. J. P., Glycoproteomics and glycomics investigation of membrane N‐glycosylproteins from human colon carcinoma cells. **2008,** *8* (16), 3236-3256.
5. Vukobrat-Bijedic, Z.; Husic-Selimovic, A.; Sofic, A.; Bijedic, N.; Bjelogrlic, I.; Gogov, B.; Mehmedovic, A. J. m. a., Cancer antigens (CEA and CA 19-9) as markers of advanced stage of colorectal carcinoma. **2013,** *67* (6), 397.
6. Chen, C.-C.; Yang, S.-H.; Lin, J.-K.; Lin, T.-C.; Chen, W.-S.; Jiang, J.-K.; Wang, H.-S.; Chang, S.-C. J. J. o. s. r., Is it reasonable to add preoperative serum level of CEA and CA19-9 to staging for colorectal cancer? **2005,** *124* (2), 169-174.
7. Varki, A.; Cummings, R. D.; Esko, J. D.; Freeze, H. H.; Stanley, P.; Bertozzi, C. R.; Hart, G. W.; Etzler, M. E., *Nematoda--Essentials of Glycobiology*. Cold Spring Harbor Laboratory Press: 2009.
8. Feizi, T. J. C. s., Carbohydrate antigens in human cancer. **1985,** *4* (1), 245-269.
9. Dall'olio, F.; Malagolini, N.; Guerrini, S.; Lau, J. T.; Serafini-Cessi, F. J. G. j., Differentiation-dependent expression of humanβ-galactosideα2, 6-sialyltransferase mRNA in colon carcinoma CaCo-2 cells. **1996,** *13* (1), 115-121.
10. Nakamori, S.; Kameyama, M.; Imaoka, S.; Furukawa, H.; Ishikawa, O.; Sasaki, Y.; Kabuto, T.; Iwanaga, T.; Matsushita, Y.; Irimura, T. J. C. r., Increased expression of sialyl Lewisx antigen correlates with poor survival in patients with colorectal carcinoma: clinicopathological and immunohistochemical study. **1993,** *53* (15), 3632-3637.
11. Guo, H.; Nagy, T.; Pierce, M. J. J. o. B. C., Post-translational glycoprotein modifications regulate colon cancer stem cells and colon adenoma progression in Apcmin/+ mice through altered Wnt receptor signaling. **2014,** *289* (45), 31534-31549.

12. Vasaikar, S.; Huang, C.; Wang, X.; Petyuk, V. A.; Savage, S. R.; Wen, B.; Dou, Y.; Zhang, Y.; Shi, Z.; Arshad, O. A. J. C., Proteogenomic analysis of human colon cancer reveals new therapeutic opportunities. **2019,** *177* (4), 1035-1049. e19.
13. Jiang, Y.; Sun, A.; Zhao, Y.; Ying, W.; Sun, H.; Yang, X.; Xing, B.; Sun, W.; Ren, L.; Hu, B. J. N., Proteomics identifies new therapeutic targets of early-stage hepatocellular carcinoma. **2019,** *567* (7747), 257-261.
14. Ahn, J.-M.; Sung, H.-J.; Yoon, Y.-H.; Kim, B.-G.; Yang, W. S.; Lee, C.; Park, H.-M.; Kim, B.-J.; Kim, B.-G.; Lee, S.-Y. J. M.; Proteomics, C., Integrated glycoproteomics demonstrates fucosylated serum paraoxonase 1 alterations in small cell lung cancer. **2014,** *13* (1), 30-48.
15. Shah, P.; Wang, X.; Yang, W.; Eshghi, S. T.; Sun, S.; Hoti, N.; Chen, L.; Yang, S.; Pasay, J.; Rubin, A. J. M.; Proteomics, C., Integrated proteomic and glycoproteomic analyses of prostate cancer cells reveal glycoprotein alteration in protein abundance and glycosylation. **2015,** *14* (10), 2753-2763.
16. Campos, D.; Freitas, D.; Gomes, J.; Magalhães, A.; Steentoft, C.; Gomes, C.; Vester-Christensen, M. B.; Ferreira, J. A.; Afonso, L. P.; Santos, L. L. J. M.; Proteomics, C., Probing the O-glycoproteome of gastric cancer cell lines for biomarker discovery. **2015,** *14* (6), 1616-1629.
17. Medzihradszky, K. F. J. M. i. e., In‐solution digestion of proteins for mass spectrometry. **2005,** *405*, 50-65.
18. Jiang, K.; Zhu, H.; Xiao, C.; Liu, D.; Edmunds, G.; Wen, L.; Ma, C.; Li, J.; Wang, P. G. J. A. c. a., Solid-phase reductive amination for glycomic analysis. **2017,** *962*, 32-40.
19. Küster, B.; Mann, M. J. A. c., 18O-labeling of N-glycosylation sites to improve the identification of gel-separated glycoproteins using peptide mass mapping and database searching. **1999,** *71* (7), 1431-1440.
20. Chi, H.; Liu, C.; Yang, H.; Zeng, W.-F.; Wu, L.; Zhou, W.-J.; Wang, R.-M.; Niu, X.-N.; Ding, Y.-H.; Zhang, Y. J. N. b., Comprehensive identification of peptides in tandem mass spectra using an efficient open search engine. **2018,** *36* (11), 1059-1061.
21. Zhu, H.; Li, X.; Qu, J.; Xiao, C.; Jiang, K.; Gashash, E.; Liu, D.; Song, J.; Cheng, J.; Ma, C. J. A.; chemistry, b., Diethylaminoethyl Sepharose (DEAE-Sepharose) microcolumn for enrichment of glycopeptides. **2017,** *409* (2), 511-518.
22. Jeong, K.; Kim, S.; Bandeira, N. J. B. b., False discovery rates in spectral identification. **2012,** *13* (S16), S2.
23. Toghi Eshghi, S.; Shah, P.; Yang, W.; Li, X.; Zhang, H. J. A. c., GPQuest: a spectral library matching algorithm for site-specific assignment of tandem mass spectra to intact N-glycopeptides. **2015,** *87* (10), 5181-5188.
24. Chambers, M. C.; Maclean, B.; Burke, R.; Amodei, D.; Ruderman, D. L.; Neumann, S.; Gatto, L.; Fischer, B.; Pratt, B.; Egertson, J. J. N. b., A cross-platform toolkit for mass spectrometry and proteomics. **2012,** *30* (10), 918-920.
25. Yu, J.; Schorlemer, M.; Gomez Toledo, A.; Pett, C.; Sihlbom, C.; Larson, G.; Westerlind, U.; Nilsson, J. J. C. A. E. J., Distinctive MS/MS Fragmentation Pathways of Glycopeptide‐Generated Oxonium Ions Provide Evidence of the Glycan Structure. **2016,** *22* (3), 1114-1124.
26. Smith, P. e.; Krohn, R. I.; Hermanson, G.; Mallia, A.; Gartner, F.; Provenzano, M.; Fujimoto, E.; Goeke, N.; Olson, B.; Klenk, D. J. A. b., Measurement of protein using bicinchoninic acid. **1985,** *150* (1), 76-85.
27. Liu, M.-Q.; Zeng, W.-F.; Fang, P.; Cao, W.-Q.; Liu, C.; Yan, G.-Q.; Zhang, Y.; Peng, C.; Wu, J.-Q.; Zhang, X.-J. J. N. c., pGlyco 2.0 enables precision N-glycoproteomics with

comprehensive quality control and one-step mass spectrometry for intact glycopeptide identification. **2017,** *8* (1), 1-14.
28. Mi, H.; Muruganujan, A.; Huang, X.; Ebert, D.; Mills, C.; Guo, X.; Thomas, P. D. J. N. p., Protocol Update for large-scale genome and gene function analysis with the PANTHER classification system (v. 14.0). **2019,** *14* (3), 703-721.
29. Afshar-Kharghan, V. J. T. J. o. c. i., The role of the complement system in cancer. **2017,** *127* (3), 780-789.
30. Ying, L.; Zhang, F.; Pan, X.; Chen, K.; Zhang, N.; Jin, J.; Wu, J.; Feng, J.; Yu, H.; Jin, H. J. O., Complement component 7 (C7), a potential tumor suppressor, is correlated with tumor progression and prognosis. **2016,** *7* (52), 86536.
31. Sanjuan, X.; Fernandez, P. L.; Castells, A.; Castronovo, V.; Van den Brule, F.; Liu, F.-T.; Cardesa, A.; Campo, E. J. G., Differential expression of galectin 3 and galectin 1 in colorectal cancer progression. **1997,** *113* (6), 1906-1915.
32. Kim, Y.-S.; Jung, J.-A.; Kim, H.-J.; Ahn, Y. H.; Yoo, J. S.; Oh, S.; Cho, C.; Yoo, H.-S.; Ko, J.-H. J. B.; communications, b. r., Galectin-3 binding protein promotes cell motility in colon cancer by stimulating the shedding of protein tyrosine phosphatase kappa by proprotein convertase 5. **2011,** *404* (1), 96-102.
33. Wojtukiewicz, M. Z.; Zacharski, L. R.; Memoli, V. A.; Kisiel, W.; Kudryk, B. J.; Rousseau, S. M.; Stump, D. C. J. T.; haemostasis, Indirect activation of blood coagulation in colon cancer. **1989,** *62* (04), 1062-1066.
34. Duke-Cohan, J. S.; Gu, J.; McLaughlin, D. F.; Xu, Y.; Freeman, G. J.; Schlossman, S. F. J. P. o. t. N. A. o. S., Attractin (DPPT-L), a member of the CUB family of cell adhesion and guidance proteins, is secreted by activated human T lymphocytes and modulates immune cell interactions. **1998,** *95* (19), 11336-11341.
35. Zhu, H.-F.; Zhang, X.-H.; Gu, C.-S.; Zhong, Y.; Long, T.; Ma, Y.-D.; Hu, Z.-Y.; Li, Z.-G.; Wang, X.-Y. J. C. b.; therapy, Cancer-associated fibroblasts promote colorectal cancer progression by secreting CLEC3B. **2019,** *20* (7), 967-978.
36. Xiao, W.; Wang, J.; Li, H.; Xia, D.; Yu, G.; Yao, W.; Yang, Y.; Xiao, H.; Lang, B.; Ma, X. J. B. c., Fibulin-1 is epigenetically down-regulated and related with bladder cancer recurrence. **2014,** *14* (1), 677.
37. Pupa, S. M.; Giuffré, S.; Castiglioni, F.; Bertola, L.; Cantú, M.; Bongarzone, I.; Baldassari, P.; Mortarini, R.; Argraves, W. S.; Anichini, A. J. C. r., Regulation of breast cancer response to chemotherapy by fibulin-1. **2007,** *67* (9), 4271-4277.
38. Levels, J.; Marquart, J.; Abraham, P.; Van den Ende, A.; Molhuizen, H.; Van Deventer, S.; Meijers, J. J. I.; immunity, Lipopolysaccharide is transferred from high-density to low-density lipoproteins by lipopolysaccharide-binding protein and phospholipid transfer protein. **2005,** *73* (4), 2321-2326.
39. Hsu, R. Y.; Chan, C. H.; Spicer, J. D.; Rousseau, M. C.; Giannias, B.; Rousseau, S.; Ferri, L. E. J. C. r., LPS-induced TLR4 signaling in human colorectal cancer cells increases β1 integrin-mediated cell adhesion and liver metastasis. **2011,** *71* (5), 1989-1998.
40. Zhu, W.; Smith, J. W.; Huang, C.-M. J. B. R. I., Mass spectrometry-based label-free quantitative proteomics. **2009,** *2010*.
41. Chelius, D.; Bondarenko, P. V. J. J. o. p. r., Quantitative profiling of proteins in complex mixtures using liquid chromatography and mass spectrometry. **2002,** *1* (4), 317-323.
42. Becker, D. J.; Lowe, J. B. J. G., Fucose: biosynthesis and biological function in mammals. **2003,** *13* (7), 41R-53R.

43. Osuga, T.; Takimoto, R.; Ono, M.; Hirakawa, M.; Yoshida, M.; Okagawa, Y.; Uemura, N.; Arihara, Y.; Sato, Y.; Tamura, F. J. J. o. t. N. C. I., Relationship between increased fucosylation and metastatic potential in colorectal cancer. **2016,** *108* (8), djw038.
44. Haltiwanger, R. S. J. G., Fucose is on the TRAIL of colon cancer. **2009,** *137* (1), 36-39.
45. Zhang, Z.; Wuhrer, M.; Holst, S. J. G. j., Serum sialylation changes in cancer. **2018,** *35* (2), 139-160.
46. Bresalier, R. S.; Ho, S. B.; Schoeppner, H. L.; Kim, Y. S.; Sleisenger, M. H.; Brodt, P.; Byrd, J. C. J. G., Enhanced sialylation of mucin-associated carbohydrate structures in human colon cancer metastasis. **1996,** *110* (5), 1354-1367.
47. Ladd, A. G.-Z.; Vásquez, A. A.; Siemes, C.; Yazdanpanah, M.; Coebergh, J. W.; Hofman, A.; Stricker, B. C.; van Duijn, C. J. B. c. r.; treatment, Differential roles of Angiotensinogen and Angiotensin Receptor type 1 polymorphisms in breast cancer risk. **2007,** *101* (3), 299-304.
48. Gimenez-Roqueplo, A.-P.; Célérier, J.; Lucarelli, G.; Corvol, P.; Jeunemaitre, X. J. J. o. B. C., Role of N-glycosylation in human angiotensinogen. **1998,** *273* (33), 21232-21238.
49. Danø, K.; Behrendt, N.; Høyer-Hansen, G.; Johnsen, M.; Lund, L. R.; Ploug, M.; Rømer, J. J. T.; haemostasis, Plasminogen activation and cancer. **2005,** *93* (04), 676-681.
50. Berger, D. H. J. W. j. o. s., Plasmin/plasminogen system in colorectal cancer. **2002,** *26* (7), 767.

3.7 Appendix

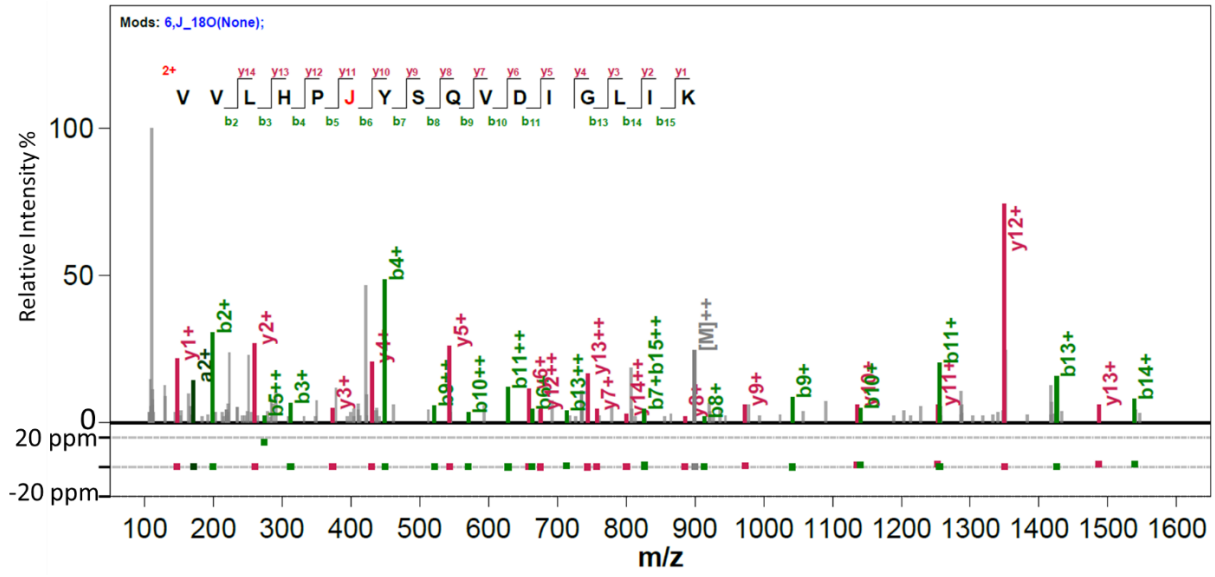


Figure 3.10 Tandem MS annotation of 18O labelled N-glycosylation site (J) of haptoglobin.

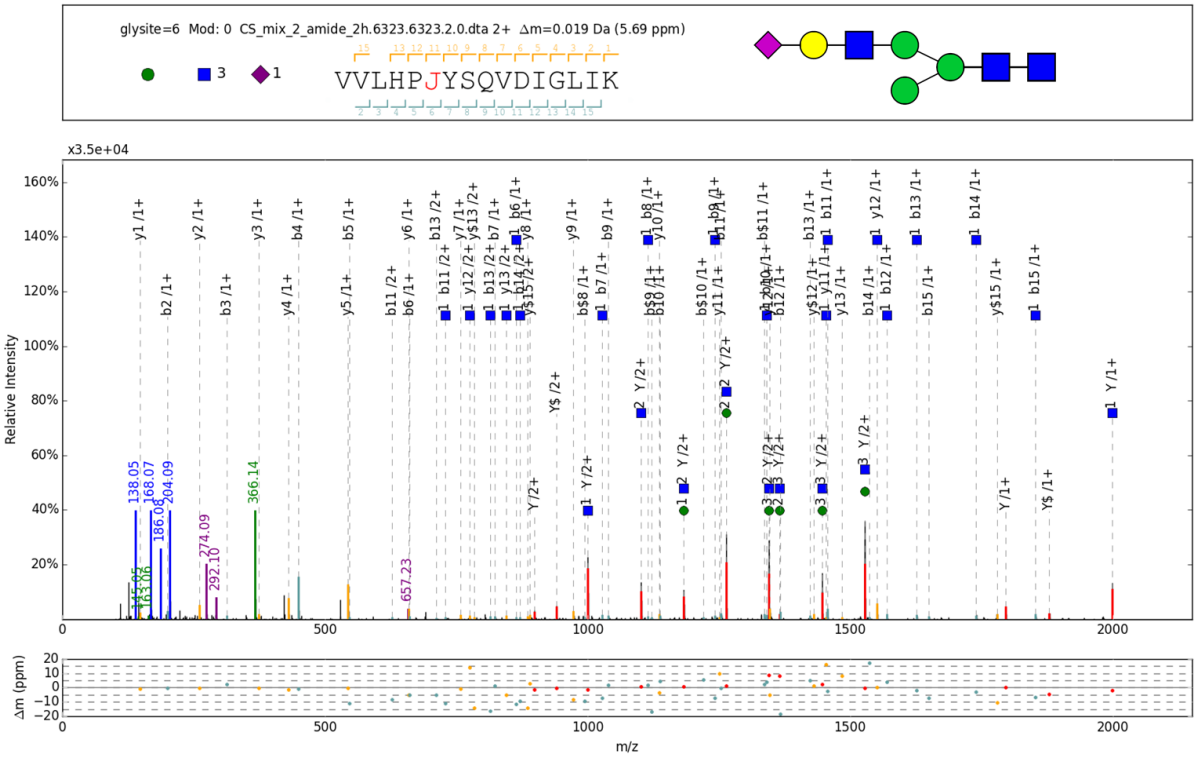


Figure 3.11 Tandem MS annotation of intact-glycopeptide from glycoprotein haptoglobin.

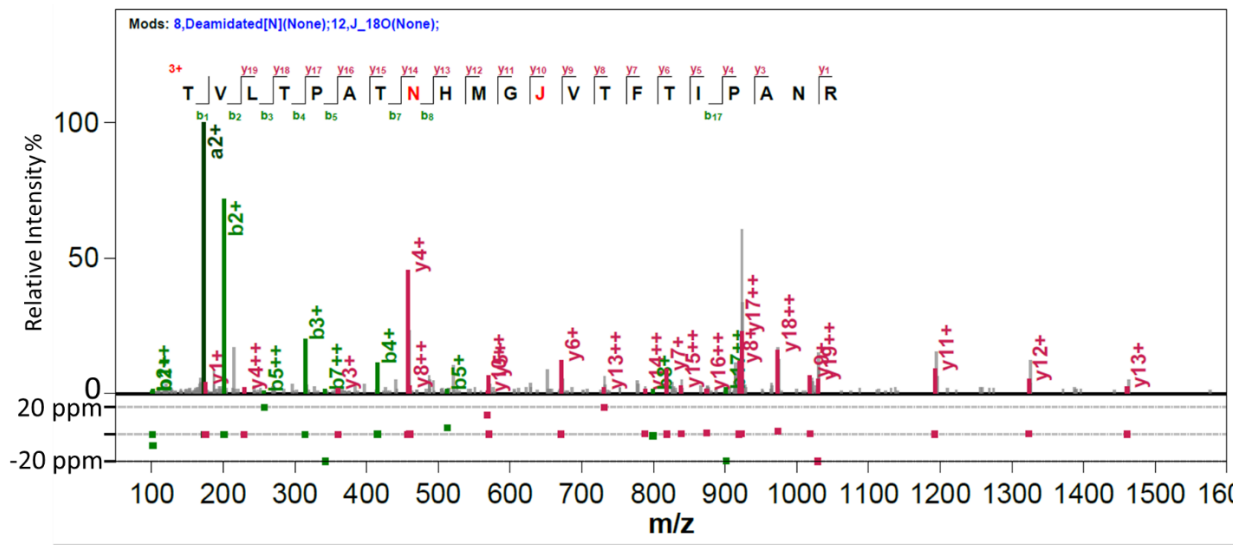


Figure 3.12 Tandem MS annotation of 18O labelled N-glycosylation site (J) of Complement C3.

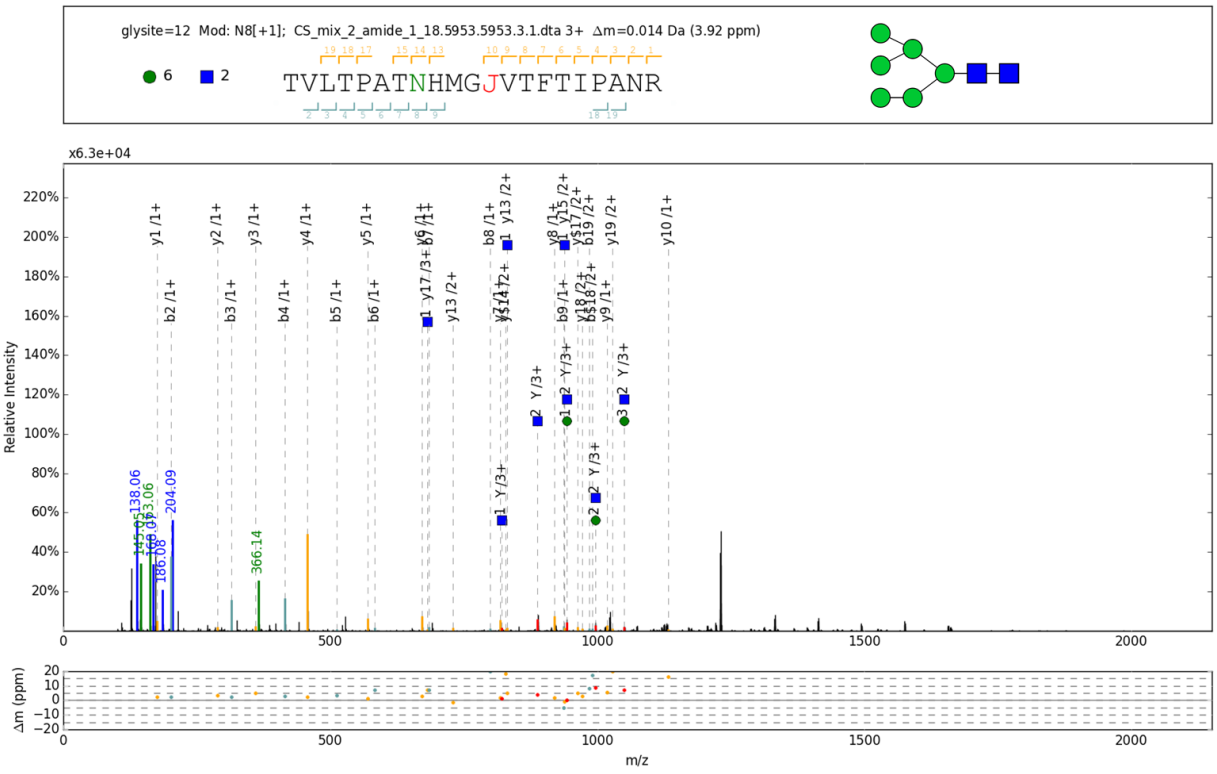


Figure 3.13 Tandem MS annotation of intact-glycopeptide from glycoprotein Haptoglobin.

Table 3.4 Protein concentration of human serum samples.

Sample	Mean of Absorbance	SEM	RSD%	Diluted Concentration (mg/mL)	Sample Concentration (mg/mL)
colon cancer 1	1.381	0.027	0.020	1.244	62.226
colon cancer 2	1.506	0.036	0.024	1.370	68.505
colon cancer 3	1.230	0.033	0.027	1.093	54.689
colon cancer 4	1.439	0.017	0.012	1.302	65.141
colon cancer 5	1.642	0.017	0.010	1.506	75.308
colon cancer 6	1.519	0.027	0.018	1.383	69.174
colon cancer 7	1.525	0.119	0.078	1.389	69.449
colon cancer 8	1.252	0.023	0.018	1.115	55.785
colon cancer 9	1.385	0.026	0.019	1.249	62.453
colon cancer 10	1.462	0.022	0.015	1.325	66.296
normal serum 1	1.680	0.007	0.004	1.544	77.220
normal serum 2	1.817	0.000	0.000	1.682	84.099
normal serum 3	1.428	0.037	0.026	1.291	64.582
normal serum 4	1.519	0.004	0.002	1.383	69.182
normal serum 5	1.659	0.013	0.007	1.523	76.170
normal serum 6	1.814	0.024	0.013	1.679	83.972
normal serum 7	1.644	0.058	0.035	1.508	75.439
normal serum 8	1.871	0.023	0.012	1.736	86.821
normal serum 9	1.755	0.052	0.029	1.619	80.973
normal serum 10	1.623	0.020	0.012	1.487	74.357

Table 3.5 Unique Glycoprotein and unique N-glycosylation sites for colon cancer serum (CS) and healthy human serum (NS).

CS unique Glycoprotein	**CS**	**NS**
Immunoglobulin lambda variable 3-16	3	0
NS unique Glycoprotein		
Homeobox protein Hox-A3	0	3
Homeobox protein Hox-B3	0	3
Homeobox protein Hox-D3	0	3
Carbonic anhydrase 1	0	3
CS unique Glycosylation site		
none		
NS unique Glycosylation site		
Carboxypeptidase B2 (N73)	0	3
Plasminogen (N308)	0	3

Table 3.6 Quantitatively analysis of identified proteins.

Protein ID	PSM of Colon Cancer	RSD	PSM of Healthy Human	RSD
P69891	0.026789	0.015327	0.048098	0.216075
P68871	0.236723	0.177776	0.453061	0.110979
P02042	0.107278	0.337899	0.245187	0.056436
P69905	0.094001	0.259916	0.276773	0.120617
P0DJI8	0.013513	1	0.075838	0.056582
P0DJI9	0.004504	1.732050	0.093646	0.311735
A0A0C4D	0.053698	0.261652	0.010647	1.021470
A0A0G2J	0.035680	0.210564	0.020715	0.043901
A0A0J9Y	0.062470	0.117539	0.038030	0.175002
O95445	0.049193	0.171118	0.010136	1.006778
P00740	0.053461	0.236337	0.027450	0.180529
P01019	0.142840	0.048680	0.086776	0.228795
P01718	0.044570	0.341255	0.013759	0.419683
P01880	0.089141	0.156621	0.051567	0.174069
P02652	0.218942	0.150068	0.137900	0.163727
P02747	0.057965	0.116886	0.038030	0.175002
P02751	1.844001	0.184381	0.841733	0.165377
P04275	0.201991	0.825980	0.034406	0.143735
P06276	0.040066	0.319690	0.003401	1.732050
P06396	0.401849	0.168838	0.245238	0.175564
P06727	0.478544	0.177736	0.306447	0.090172
P07996	0.143077	0.156533	0.045634	0.621289
P08185	0.102892	0.211969	0.062215	0.177210
P23142	0.085111	0.339889	0.031005	0.319097

Table 3.7 Quantitatively analysis of identified glycoproteins.

Glycoprotein ID	PSM of Colon Cancer	RSD	PSM of Healthy Human	RSD
P01011	0.8533	0.1083	0.5067	0.1987
P02647	0.2000	0.3000	0.0267	0.0375
P03952	0.3200	0.1250	0.1733	0.3525
P02652	0.2000	0.2000	0.0667	0.6928
P01019	0.1467	0.1575	0.0400	0.0125
P02748	0.1867	0.3000	0.0400	0.0250
P07225	0.1200	0.0000	0.0133	0.0750
P12259	0.0933	0.4949	0.0133	0.0375
P05452	0.0667	0.3464	0.0133	0.0750
P80108	0.0933	0.2474	0.1467	0.1575

Table 3.8 All Identified intact-glycopeptide from cancer and healthy human serum.

Precursor m/z	Charge	Protein Accession	Peptide	NGLYCAN
1127.301392	4	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N2H7F0S0G0:4H
1097.043335	4	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N3H5F0S0G0:4H
1143.748901	4	sp|P02765;	AALAAFNAQNNGSNFQLEEISR	N4H5F4S0G0:4H
1526.74707	2	sp|P00738;	VVLHPNYSQVDIGLIK	N3H4F0S0G0:2H
1263.148804	2	sp|P00738;	VVLHPNYSQVDIGLIK	N2H2F0S0G0:2H
1155.736572	4	sp|P01871;	GLTFQQNASSM[Oxidation]C[Carbamidomethyl]VPDQDTAIR	N5H5F3S0G0:4H
1354.015503	3	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N3H3F0S0G0:3H
1820.2455	2	sp|P0C0L4;sp|P0C0L5;	FSDGLESNSSTQFEVK	N2H9F0S0G0:2H
1394.356567	3	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N2H5F0S0G0:3H
1591.77063	2	sp|P00738;	VVLHPNYSQVDIGLIK	N3H3F2S0G0:2H
1275.348755	4	sp|P01009;	QLAHQSNSTNIFFSPVSIATAFAMLSLGTK	N4H5F2S0G0:4H
1526.7385	2	sp|P00738;	VVLHPNYSQVDIGLIK	N3H4F0S0G0:2H
1283.079102	2	sp|P01877;	TPLTANITK	N4H4F1S0G0:2H
1286.647827	3	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N2H3F0S0G0:3H
1344.175171	2	sp|P00738;	VVLHPNYSQVDIGLIK	N2H3F0S0G0:2H
1259.277222	3	sp|P00738;sp|P00739;	MVSHHNLTTGATLINEQWLLTTAK	N3H3F0S0G0:3H
1445.718262	2	sp|P00738;	VVLHPNYSQVDIGLIK	N3H3F0S0G0:2H
1129.300903	4	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N3H4F2S0G0:4H
1672.2891	2	sp|P00738;	VVLHPNYSQVDIGLIK	N3H4F0S1G0:2H
1106.797241	4	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N4H4F0S0G0:4H
1384.635864	4	sp|P01009;	QLAHQSNSTNIFFSPVSIATAFAMLSLGTK	N6H7F0S0G0:4H
1017.4328	2	sp|P01859;	EEQFNSTFR	N2H2F1S0G0:2H
1268.0522	2	sp|O75882;	IDSTGNVTNELR	N2H5F0S0G0:2H
1181.546387	2	sp|P01877;	TPLTANITK	N3H4F1S0G0:2H
1214.159912	3	sp|P0C0L4;sp|P0C0L5;	FSDGLESNSSTQFEVK	N2H9F0S0G0:3H
1263.146606	2	sp|P00738;	VVLHPNYSQVDIGLIK	N2H2F0S0G0:2H
1344.171143	2	sp|P00738;	VVLHPNYSQVDIGLIK	N2H3F0S0G0:2H
1093.253906	4	sp|P01023;	VSNQTLSLFFTVLQDVPVR	N4H5F4S0G0:4H
1100.971313	4	sp|P01871;	GLTFQQNASSMC[Carbamidomethyl]VPDQDTAIR	N4H5F3S0G0:4H
1018.164795	3	sp|P00738;	VVLHPNYSQVDIGLIK	N3H4F0S0G0:3H
1312.960083	3	sp|P00738;sp|P00739;	MVSHHNLTTGATLINEQWLLTTAK	N3H4F0S0G0:3H
1410.331909	3	sp|P00738;sp|P00739;	MVSHHNLTTGATLINEQWLLTTAK	N3H4F2S0G0:3H
847.9077759	6	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N5H5F2S0G0:6H
1475.668579	4	sp|P01009;	ADTHDEILEGLNFNLTEIPEAQIHEGFQELLR	N4H5F4S0G0:4H
1107.297363	4	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N4H4F0S0G0:4H
984.9747925	4	sp|P00738;sp|P00739;	MVSHHNLTTGATLINEQWLLTTAK	N3H4F0S0G0:4H
1434.6742	3	sp|P00738;sp|P00739;	MVSHHNLTTGATLINEQWLLTTAK	N4H5F0S0G0:3H
1139.879883	3	sp|P00738;	VVLHPNYSQVDIGLIK	N4H5F0S0G0:3H
1434.671753	3	sp|P00738;sp|P00739;	MVSHHNLTTGATLINEQWLLTTAK	N4H5F0S0G0:3H
1078.697632	5	sp|P01009;	QLAHQSNSTNIFFSPVSIATAFAMLSLGTK	N4H5F4S0G0:5H
1081.899048	5	sp|P01009;	QLAHQSNSTNIFFSPVSIATAFAM[Oxidation]LSLGTK	N4H5F4S0G0:5H

996.152771	3	sp|P01024;	TVLTPATNHMGNVTFTIPANR	N2H2F0S0G0:3H
1158.21106	3	sp|P01024;	TVLTPATNHMGNVTFTIPANR	N2H5F0S0G0:3H
1104.187134	3	sp|P01024;	TVLTPATNHMGNVTFTIPANR	N2H4F0S0G0:3H
949.6871948	4	sp|P01024;	TVLTPATNHMGNVTFTIPANR	N2H7F0S0G0:4H
1202.055542	2	sp|P01877;	TPLTANITK	N4H3F1S0G0:2H
953.6860352	4	sp|P01024;	TVLTPATNHM[Oxidation]GNVTFTIPANR	N2H7F0S0G0:4H
1224.20813	3	sp|P01009;	YLGNATAIFFLPDEGK	N4H5F2S0G0:3H
813.7265015	6	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N4H4F1S0G1:6H
1318.624512	3	sp|P00738;sp|P00739;	M[Oxidation]VSHHNLTTGATLINEQWLLTTAK	N3H4F0S0G0:3H
1817.3367	2	sp|P01024;	TVLTPATNHMGNVTFTIPANR	N2H6F0S0G0:2H
1445.713745	2	sp|P00738;	VVLHPNYSQVDIGLIK	N3H3F0S0G0:2H
1018.763916	3	sp|P01871;	YKNNSDISSTR	N4H5F1S0G0:3H
1407.696045	3	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N3H4F0S0G0:3H
1117.55603	4	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N5H3F0S0G0:4H
1422.040283	3	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N4H3F0S0G0:3H
1532.038818	3	sp|P00738;sp|P00739;	MVSHHNLTTGATLINEQWLLTTAK	N4H5F2S0G0:3H
1140.215332	3	sp|P00738;	VVLHPNYSQVDIGLIK	N4H5F0S0G0:3H
1205.0385	4	sp|P00450;	ELHHLQEQNVSNAFLDKGEFYIGSK	N4H4F1S0G1:4H
1388.631348	4	sp|P01009;	QLAHQSNSTNIFFSPVSIATAFAM[Oxidation]LSLGTK	N6H7F0S0G0:4H
1492.705444	4	sp|P01023;	TEVSSNHVLIYLDKVSNQTLSLFFTVLQDVPVR	N4H5F2S1G0:4H
1279.344482	4	sp|P01009;	QLAHQSNSTNIFFSPVSIATAFAM[Oxidation]LSLGTK	N4H5F2S0G0:4H
1047.46875	2	sp|Q92954;	NGTLVAFR	N2H5F0S0G0:2H
960.3978882	2	sp|P01857;	EEQYNSTYR	N2H2F0S0G0:2H
1403.079468	2	sp|P01859;	EEQFNSTFR	N5H3F1S0G0:2H
1301.538452	2	sp|P01859;	EEQFNSTFR	N4H3F1S0G0:2H
1477.621826	2	sp|P02790;	SWPAVGNC[Carbamidomethyl]SSALR	N3H4F2S0G0:2H
1212.226318	3	sp|P01024;	TVLTPATNHMGNVTFTIPANR	N2H6F0S0G0:3H
1212.228027	3	sp|P01024;	TVLTPATNHMGNVTFTIPANR	N2H6F0S0G0:3H
1217.560181	3	sp|P01024;	TVLTPATNHM[Oxidation]GNVTFTIPANR	N2H6F0S0G0:3H
909.1715698	4	sp|P01024;	TVLTPATNHMGNVTFTIPANR	N2H6F0S0G0:4H
1102.161499	3	sp|P01009;	YLGNATAIFFLPDEGK	N3H4F0S1G0:3H
1056.493896	4	sp|P01023;	VSNQTLSLFFTVLQDVPVR	N4H4F2S0G1:4H
993.7540894	3	sp|P01871;	YKNNSDISSTR	N3H4F1S1G0:3H
1019.980164	4	sp|P01023;	VSNQTLSLFFTVLQDVPVR	N4H4F1S0G1:4H
964.4238892	3	sp|P27169;	HANWTLTPLK	N3H4F1S0G1:3H
1080.2549	4	sp|P00738;sp|P00739;	M[Oxidation]VSHHNLTTGATLINEQWLLTTAK	N4H5F0S0G0:4H
1415.993896	3	sp|P00738;sp|P00739;	M[Oxidation]VSHHNLTTGATLINEQWLLTTAK	N3H4F2S0G0:3H
1158.066406	4	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N5H4F0S0G0:4H
1117.556885	4	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N5H3F0S0G0:4H
1085.855957	3	sp|P00738;	VVLHPNYSQVDIGLIK	N4H4F0S0G0:3H
1113.218872	4	sp|P01011;	TLNQSSDELQLSM[Oxidation]GNAM[Oxidation]FVK	N4H5F0S2G0:4H
1076.260254	4	sp|P00738;sp|P00739;	MVSHHNLTTGATLINEQWLLTTAK	N4H5F0S0G0:4H
1220.585327	4	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N4H5F2S0G0:4H
855.378479	2	sp|P01859;	EEQFNSTFR	N2H0F1S0G0:2H

1215.99292	2	sp|P01857;	EEQYNSTYR	N3H3F1S0G0:2H
1281.03418	2	sp|P01859;	EEQFNSTFR	N3H4F1S0G0:2H
1200.000732	2	sp|P01859;	EEQFNSTFR	N3H3F1S0G0:2H
1297.017944	2	sp|P01857;	EEQYNSTYR	N3H4F1S0G0:2H
1419.074829	2	sp|P01857;	EEQYNSTYR	N5H3F1S0G0:2H
1390.567871	2	sp|P01861;	EEQFNSTYR	N4H4F1S0G0:2H
1309.539063	2	sp|P01861;	EEQFNSTYR	N4H3F1S0G0:2H
1207.994507	2	sp|P01861;	EEQFNSTYR	N3H3F1S0G0:2H
1384.6215	2	sp|P01877;	TPLTANITK	N5H4F1S0G0:2H
1465.6482	2	sp|P01877;	TPLTANITK	N5H5F1S0G0:2H
1163.871338	3	sp|P01024;	TVLTPATNHM[Oxidation]GNVTFTIPANR	N2H5F0S0G0:3H
1265.910889	3	sp|P01024;	TVLTPATNHMGNVTFTIPANR	N2H7F0S0G0:3H
1217.22522	3	sp|P01024;	TVLTPATNHM[Oxidation]GNVTFTIPANR	N2H6F0S0G0:3H
872.6584473	4	sp|P01024;	TVLTPATNHM[Oxidation]GNVTFTIPANR	N2H5F0S0G0:4H
913.1724854	4	sp|P01024;	TVLTPATNHM[Oxidation]GNVTFTIPANR	N2H6F0S0G0:4H
1157.875366	3	sp|P01024;	TVLTPATNHMGNVTFTIPANR	N2H5F0S0G0:3H
1344.572754	2	sp|P01871;	YKNNSDISSTR	N3H4F1S0G0:2H
1212.226074	3	sp|P01024;	TVLTPATNHMGNVTFTIPANR	N2H6F0S0G0:3H
1173.002686	2	sp|P00734;	GHVNITR	N3H4F0S1G0:2H
759.8636475	4	sp|O75882;	VFHIHNESWVLLTPK	N2H5F0S0G0:4H
1410.658081	3	sp|P00738;sp|P00739;	MVSHHNLTTGATLINEQWLLTTAK	N3H4F2S0G0:3H
984.9747314	4	sp|P00738;sp|P00739;	MVSHHNLTTGATLINEQWLLTTAK	N3H4F0S0G0:4H
1148.77	4	sp|P04114;	QVLFLDTVYGNC[Carbamidomethyl]STHFTVK	N7H4F2S0G0:4H
868.9093018	4	sp|P01024;	TVLTPATNHMGNVTFTIPANR	N2H5F0S0G0:4H
1476.056885	3	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N4H4F0S0G0:3H
1104.469849	4	sp|P01871;	GLTFQQNASSM[Oxidation]C[Carbamidomethyl]VPDQDT AIR	N4H4F2S0G1:4H
1209.828979	4	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N3H6F0S1G0:4H
1061.175293	3	sp|P00738;	VVLHPNYSQVDIGLIK	N2H5F1S0G0:2H+ Na
1524.660767	3	sp|P02765;	AALAAFNAQNNGSNFQLEEISR	N4H5F4S0G0:3H
1061.518555	3	sp|P00738;	VVLHPNYSQVDIGLIK	N3H3F2S0G0:3H
1634.737549	3	sp|P00738;sp|P00739;	M[Oxidation]VSHHNLTTGATLINEQWLLTTAK	N4H5F4S0G0:3H
1237.243286	3	sp|P00738;	VVLHPNYSQVDIGLIK	N4H5F2S0G0:3H
1271.607788	4	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N5H5F2S0G0:4H
1532.371704	3	sp|P00738;sp|P00739;	MVSHHNLTTGATLINEQWLLTTAK	N4H5F2S0G0:3H
1537.037231	3	sp|P00738;sp|P00739;	M[Oxidation]VSHHNLTTGATLINEQWLLTTAK	N4H4F1S0G1:3H
1147.814819	4	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N4H5F0S0G0:4H
1628.7345	3	sp|P00738;sp|P00739;	MVSHHNLTTGATLINEQWLLTTAK	N4H5F0S2G0:3H
1199.998535	2	sp|P01859;	EEQFNSTFR	N3H3F1S0G0:2H
1033.42334	2	sp|P01857;	EEQYNSTYR	N2H2F1S0G0:2H
1207.994019	2	sp|P01861;	EEQFNSTYR	N3H3F1S0G0:2H
1289.020996	2	sp|P01861;	EEQFNSTYR	N3H4F1S0G0:2H
1199.996338	2	sp|P01859;	EEQFNSTFR	N3H3F1S0G0:2H
1228.514282	2	sp|P01859;	EEQFNSTFR	N4H3F0S0G0:2H
1215.9932	2	sp|P01857;	EEQYNSTYR	N3H3F1S0G0:2H

1223.99292	2	sp|P01857;	EEQYNSTYR	N3H4F0S0G0:2H
1484.1	2	sp|P01859;	EEQFNSTFR	N5H4F1S0G0:2H
1382.569	2	sp|P01859;	EEQFNSTFR	N4H4F1S0G0:2H
1215.999878	2	sp|P01857;	EEQYNSTYR	N3H3F1S0G0:2H
1281.028	2	sp|P01859;	EEQFNSTFR	N3H4F1S0G0:2H
1142.965942	2	sp|P01857;	EEQYNSTYR	N3H3F0S0G0:2H
1244.508911	2	sp|P01857;	EEQYNSTYR	N4H3F0S0G0:2H
1217.228394	3	sp|P01024;	TVLTPATNHM[Oxidation]GNVTFTIPANR	N2H6F0S0G0:3H
1217.889404	3	sp|P01024;	TVLTPATNHM[Oxidation]GNVTFTIPANR	N2H6F0S0G0:3H
1055.427002	3	sp|P02749;	LGNWSAMPSC[Carbamidomethyl]K	N4H4F1S0G1:3H
1595.0614	3	sp|P05090;	ADGTVNQIEGEATPVNLTEPAKLEVK	N4H4F2S0G1:3H
1202.843018	3	sp|P0C0L4;sp|P0C0L5;	FSDGLESNSSTQFEVKK	N2H8F0S0G0:3H
1263.545898	2	sp|P01871;	YKNNSDISSTR	N3H3F1S0G0:2H
1256.862183	3	sp|P0C0L4;sp|P0C0L5;	FSDGLESNSSTQFEVKK	N2H9F0S0G0:3H
1147.470215	3	sp|P04114;	YDFNSSMLYSTAK	N4H4F1S0G1:3H
820.8980103	6	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N5H4F2S0G0:6H
1407.696533	3	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N3H4F0S0G0:3H
1169.817871	4	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N3H4F1S0G1:4H
1057.746338	4	sp|P00738;sp|P00739;	MVSHHNLTTGATLINEQWLLTTAK	N3H4F0S1G0:4H
1376.623901	3	sp|P01024;	TVLTPATNHMGNVTFTIPANR	N3H6F0S1G0:3H
918.1599121	4	sp|P01009;	YLGNATAIFFLPDEGK	N4H4F1S0G1:4H
1058.251221	4	sp|P00738;sp|P00739;	MVSHHNLTTGATLINEQWLLTTAK	N3H4F2S0G0:4H
1537.371338	3	sp|P00738;sp|P00739;	M[Oxidation]VSHHNLTTGATLINEQWLLTTAK	N4H5F2S0G0:3H
1317.535645	2	sp|P01857;	EEQYNSTYR	N4H3F1S0G0:2H
1398.560791	2	sp|P01857;	EEQYNSTYR	N4H4F1S0G0:2H
1492.0972	2	sp|P01861;	EEQFNSTYR	N5H4F1S0G0:2H
1479.585449	2	sp|P01857;	EEQYNSTYR	N4H5F1S0G0:2H
1325.533569	2	sp|P01857;	EEQYNSTYR	N4H4F0S0G0:2H
1317.536743	2	sp|P01857;	EEQYNSTYR	N4H3F1S0G0:2H
1317.531	2	sp|P01857;	EEQYNSTYR	N4H3F1S0G0:2H
1500.0979	2	sp|P01857;	EEQYNSTYR	N5H4F1S0G0:2H
1301.534668	2	sp|P01859;	EEQFNSTFR	N4H3F1S0G0:2H
1271.243164	3	sp|P01024;	TVLTPATNHM[Oxidation]GNVTFTIPANR	N2H7F0S0G0:3H
874.7033691	3	sp|P04004;	NGSLFAFR	N3H4F1S0G1:3H
1142.968018	2	sp|P01857;	EEQYNSTYR	N3H3F0S0G0:2H
1611.2111	2	sp|P01877;	TPLTANITK	N5H5F1S1G0:2H
1364.1104	2	sp|P01877;	TPLTANITK	N4H5F1S0G0:2H
1509.6566	2	sp|P01877;	TPLTANITK	N4H4F2S0G1:2H
981.3981323	3	sp|P01861;	EEQFNSTYR	N4H5F1S0G0:3H
945.0963745	3	sp|P06681;	QSVPAHFVALNGSK	N2H6F0S0G0:3H
1130.82666	3	sp|P02787;	C[Carbamidomethyl]GLVPVLAENYNK	N4H4F1S0G1:3H
1024.74231	3	sp|P05546;	DFVNASSK	N4H5F0S2G0:3H
1182.8927	3	sp|P00738;	VVLHPNYSQVDIGLIK	N4H4F0S1G0:3H
1529.744385	3	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N4H5F0S0G0:3H

1080.983276	2	sp|P01871;	YKNNSDISSTR	N2H2F1S0G0:2H
1489.734375	3	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N5H3F0S0G0:3H
1148.14209	3	sp|P00450;	EHEGAIYPDNTTDFQR	N3H4F2S0G0:3H
1216.147461	3	sp|P08603;	M[Oxidation]DGASNVTC[Carbamidomethyl]INSR	N4H5F2S1G0:3H
1062.249756	4	sp|P00738;sp|P00739;	M[Oxidation]VSHHNLTTGATLINEQWLLTTAK	N3H4F2S0G0:4H
1530.075073	3	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N4H5F0S0G0:3H
1152.4579	3	sp|P02749;	LGNWSAMPSC[Carbamidomethyl]K	N4H5F0S2G0:3H
1283.0808	2	sp|P01877;	TPLTANITK	N4H4F1S0G0:2H
1655.1973	2	sp|P0C0L4;sp|P0C0L5;	GLNVTLSSTGR	N4H5F0S2G0:2H
1321.574097	3	sp|P01009;	YLGNATAIFFLPDEGK	N4H5F4S0G0:3H
1020.45282	3	sp|P01877;	TPLTANITK	N5H4F1S1G0:3H
1472.954712	3	sp|P01871;	GLTFQQNASSM[Oxidation]C[Carbamidomethyl]VPDQDTAIR	N4H5F3S0G0:3H
1037.215576	4	sp|P01024;	TVLTPATNHM[Oxidation]GNVTFTIPANR	N3H6F0S1G0:4H
836.9037476	4	sp|P00738;	VVLHPNYSQVDIGLIK	N3H4F2S0G0:4H
1057.528076	4	sp|P05155;	VGQLQLSHNLSLVILVPQNLK	N4H4F1S0G1:4H
1009.762268	3	sp|P13671;	VLNFTTK	N4H5F0S2G0:3H
1322.937744	3	sp|P01024;	TVLTPATNHMGNVTFTIPANR	N3H4F1S0G1:3H
1158.537598	3	sp|P01024;	TVLTPATNHMGNVTFTIPANR	N2H5F0S0G0:3H
1217.55542	3	sp|P01024;	TVLTPATNHM[Oxidation]GNVTFTIPANR	N2H6F0S0G0:3H
1217.2148	3	sp|P01024;	TVLTPATNHM[Oxidation]GNVTFTIPANR	N2H6F0S0G0:3H
909.4194946	4	sp|P01024;	TVLTPATNHMGNVTFTIPANR	N2H6F0S0G0:4H
1023.757202	3	sp|P03952;	GVNFNVSK	N4H5F0S2G0:3H
1514.1373	2	sp|P13671;	VLNFTTK	N4H5F0S2G0:2H
1311.5487	2	sp|P04004;	NGSLFAFR	N3H4F1S0G1:2H
1212.556152	3	sp|P01024;	TVLTPATNHMGNVTFTIPANR	N2H6F0S0G0:3H
1158.210571	3	sp|P01024;	TVLTPATNHMGNVTFTIPANR	N2H5F0S0G0:3H
702.6040649	3	DECOY	SYDAHEEWMNGFDSNT	N1H0F0S0G0:3H
1544.0324	3	sp|P01023;	SLGNVNFTVSAEALESQELC[Carbamidomethyl]GTEVPSVPEHGR	N2H5F0S0G0:3H
1127.301758	4	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N2H7F0S0G0:4H
1015.5107	4	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N3H3F0S0G0:4H
1158.78064	4	sp|P01023;	SLGNVNFTVSAEALESQELC[Carbamidomethyl]GTEVPSVPEHGR	N2H5F0S0G0:4H
1046.272583	4	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N2H5F0S0G0:4H
1168.064453	4	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N6H3F0S0G0:4H
1181.548706	2	sp|P01877;	TPLTANITK	N3H4F1S0G0:2H
1102.497559	3	sp|P01009;	YLGNATAIFFLPDEGK	N3H4F2S0G0:3H
1394.355469	3	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N2H5F0S0G0:3H
1263.149658	2	sp|P00738;	VVLHPNYSQVDIGLIK	N2H2F0S0G0:2H
1066.782959	4	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N4H3F0S0G0:4H
903.8495483	5	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N3H4F2S0G0:5H
814.2287598	6	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N4H5F2S0G0:6H
1025.430908	2	sp|P01861;	EEQFNSTYR	N2H2F1S0G0:2H
1591.267456	2	sp|P00738;	VVLHPNYSQVDIGLIK	N2H5F1S0G0:H+Na
1422.045166	3	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N4H3F0S0G0:3H

1527.241455	2	sp|P00738;	VVLHPNYSQVDIGLIK	N3H4F0S0G0:2H
1445.723145	2	sp|P00738;	VVLHPNYSQVDIGLIK	N3H3F0S0G0:2H
847.9100952	6	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N5H5F2S0G0:6H
1356.311401	3	sp|P00738;sp|P00739;	MVSHHNLTTGATLINEQWLLTTAK	N3H3F2S0G0:3H
1353.679565	3	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N3H3F0S0G0:3H
1348.123901	4	sp|P01009;	QLAHQSNSTNIFFSPVSIATAFAMLSLGTK	N4H5F4S0G0:4H
1017.430908	2	sp|P01859;	EEQFNSTFR	N2H2F1S0G0:2H
1820.2416	2	sp|P0C0L4;sp|P0C0L5;	FSDGLESNSSTQFEVK	N2H9F0S0G0:2H
1229.62561	2	sp|P01024;	TVLTPATNHMGNVTFTIPANR	N1H0F0S0G0:2H
1129.056641	4	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N3H4F2S0G0:4H
1263.152222	2	sp|P00738;	VVLHPNYSQVDIGLIK	N2H2F0S0G0:2H
1344.179321	2	sp|P00738;	VVLHPNYSQVDIGLIK	N2H3F0S0G0:2H
1107.300781	4	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N4H4F0S0G0:4H
1381.328369	3	sp|P00738;sp|P00739;	MVSHHNLTTGATLINEQWLLTTAK	N4H4F0S0G0:3H
1475.671509	4	sp|P01009;	ADTHDEILEGLNFNLTEIPEAQIHEGFQELLR	N4H5F4S0G0:4H
1529.746826	3	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N4H5F0S0G0:3H
1143.748657	4	sp|P02765;	AALAAFNAQNNGSNFQLEEISR	N4H5F4S0G0:4H
1476.064087	3	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N4H4F0S0G0:3H
1180.740479	5	sp|P01009;	ADTHDEILEGLNFNLTEIPEAQIHEGFQELLR	N4H5F4S0G0:5H
1268.057739	2	sp|O75882;	IDSTGNVTNELR	N2H5F0S0G0:2H
1283.090454	2	sp|P01877;	TPLTANITK	N4H4F1S0G0:2H
1214.501343	3	sp|P0C0L4;sp|P0C0L5;	FSDGLESNSSTQFEVK	N6H4F0S0G0:3H
1344.179077	2	sp|P00738;	VVLHPNYSQVDIGLIK	N2H3F0S0G0:2H
820.565979	6	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N5H4F0S1G0:6H
1194.084839	4	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N5H4F1S0G0:4H
1155.743408	4	sp|P01871;	GLTFQQNASSM[Oxidation]C[Carbamidomethyl]VPDQDTAIR	N5H5F3S0G0:4H
1526.749634	2	sp|P00738;	VVLHPNYSQVDIGLIK	N3H4F0S0G0:2H
1259.609863	3	sp|P00738;sp|P00739;	MVSHHNLTTGATLINEQWLLTTAK	N3H3F0S0G0:3H
1672.2886	2	sp|P00738;	VVLHPNYSQVDIGLIK	N3H4F0S1G0:2H
1313.297363	3	sp|P00738;sp|P00739;	MVSHHNLTTGATLINEQWLLTTAK	N3H4F0S0G0:3H
1318.286865	3	sp|P00738;sp|P00739;	M[Oxidation]VSHHNLTTGATLINEQWLLTTAK	N3H4F0S0G0:3H
1117.555786	4	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N5H3F0S0G0:4H
1361.642822	3	sp|P00738;sp|P00739;	M[Oxidation]VSHHNLTTGATLINEQWLLTTAK	N3H3F2S0G0:3H
1410.330933	3	sp|P00738;sp|P00739;	MVSHHNLTTGATLINEQWLLTTAK	N3H4F2S0G0:3H
894.246582	5	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N5H3F0S0G0:5H
1435.341309	3	sp|P00738;sp|P00739;	MVSHHNLTTGATLINEQWLLTTAK	N4H5F0S0G0:3H
1275.352051	4	sp|P01009;	QLAHQSNSTNIFFSPVSIATAFAMLSLGTK	N4H5F2S0G0:4H
1796.1541	3	sp|P01009;	QLAHQSNSTNIFFSPVSIATAFAMLSLGTK	N4H5F0S2G0:3H
1139.881348	3	sp|P00738;	VVLHPNYSQVDIGLIK	N4H5F0S0G0:3H
1279.098022	4	sp|P01009;	QLAHQSNSTNIFFSPVSIATAFAM[Oxidation]LSLGTK	N4H5F2S0G0:4H
1402.896118	4	sp|P01009;	ADTHDEILEGLNFNLTEIPEAQIHEGFQELLR	N4H5F2S0G0:4H
1047.473755	2	sp|Q92954;	NGTLVAFR	N2H5F0S0G0:2H
1215.995483	2	sp|P01857;	EEQYNSTYR	N3H3F1S0G0:2H

1384.6243	2	sp|P01877;	TPLTANITK	N5H4F1S0G0:2H
1477.131104	2	sp|P02790;	SWPAVGNC[Carbamidomethyl]SSALR	N3H4F0S1G0:2H
913.4237061	4	sp|P01024;	TVLTPATNHM[Oxidation]GNVTFTIPANR	N2H6F0S0G0:4H
1113.784668	3	sp|P08603;	MDGASNVTC[Carbamidomethyl]INSR	N4H5F2S0G0:3H
1056.996826	4	sp|P01023;	VSNQTLSLFFTVLQDVPVR	N4H5F3S0G0:4H
1190.804443	4	sp|P01023;	SLGNVNFTVSAEALESQELC[Carbamidomethyl]GTEVPS VPEHGRK	N2H5F0S0G0:4H
1018.167236	3	sp|P00738;	VVLHPNYSQVDIGLIK	N3H4F0S0G0:3H
1817.3308	2	sp|P01024;	TVLTPATNHMGNVTFTIPANR	N2H6F0S0G0:2H
1093.257324	4	sp|P01023;	VSNQTLSLFFTVLQDVPVR	N4H5F4S0G0:4H
1018.433411	3	sp|P01871;	YKNNSDISSTR	N4H5F1S0G0:3H
1018.169128	3	sp|P00738;	VVLHPNYSQVDIGLIK	N3H4F0S0G0:3H
1107.912476	5	sp|P01009;	QLAHQSNSTNIFFSPVSIATAFAMLSLGTK	N6H7F0S0G0:5H
1410.664185	3	sp|P00738;sp|P00739;	MVSHHNLTTGATLINEQWLLTTAK	N3H4F2S0G0:3H
1532.042725	3	sp|P00738;sp|P00739;	MVSHHNLTTGATLINEQWLLTTAK	N4H5F2S0G0:3H
1205.0394	4	sp|P00450;	ELHHLQEQNVSNAFLDKGEFYIGSK	N4H4F1S0G1:4H
1140.214966	3	sp|P00738;	VVLHPNYSQVDIGLIK	N4H5F0S0G0:3H
1147.563965	4	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N4H5F0S0G0:4H
1114.258789	4	sp|P00450;	ELHHLQEQNVSNAFLDKGEFYIGSK	N3H4F2S0G0:4H
1245.288818	4	sp|P04004;	NISDGFDGIPDNVDAALALPAHSYSGR	N4H5F2S1G0:4H
1081.900391	5	sp|P01009;	QLAHQSNSTNIFFSPVSIATAFAM[Oxidation]LSLGTK	N4H5F4S0G0:5H
1194.776611	5	sp|P01023;	TEVSSNHVLIYLDKVSNQTLSLFFTVLQDVPVR	N4H5F4S0G0:5H
1033.43103	2	sp|P01857;	EEQYNSTYR	N2H2F1S0G0:2H
1215.999146	2	sp|P01857;	EEQYNSTYR	N3H3F1S0G0:2H
1283.084	2	sp|P01877;	TPLTANITK	N4H4F1S0G0:2H
1207.999268	2	sp|P01861;	EEQFNSTYR	N3H3F1S0G0:2H
1465.6523	2	sp|P01877;	TPLTANITK	N5H5F1S0G0:2H
1158.208496	3	sp|P01024;	TVLTPATNHMGNVTFTIPANR	N2H5F0S0G0:3H
1158.210205	3	sp|P01024;	TVLTPATNHMGNVTFTIPANR	N2H5F0S0G0:3H
1142.967041	2	sp|P01857;	EEQYNSTYR	N3H3F0S0G0:2H
1163.207886	3	sp|P01024;	TVLTPATNHM[Oxidation]GNVTFTIPANR	N2H5F0S0G0:3H
1212.564819	3	sp|P01024;	TVLTPATNHMGNVTFTIPANR	N2H6F0S0G0:3H
1211.89502	3	sp|P01024;	TVLTPATNHMGNVTFTIPANR	N2H6F0S0G0:3H
1104.190308	3	sp|P01024;	TVLTPATNHMGNVTFTIPANR	N2H4F0S0G0:3H
993.7546997	3	sp|P01871;	YKNNSDISSTR	N3H4F1S1G0:3H
896.7217407	3	sp|P01871;	YKNNSDISSTR	N3H4F1S0G0:3H
1058.250854	4	sp|P00738;sp|P00739;	MVSHHNLTTGATLINEQWLLTTAK	N3H4F2S0G0:4H
1080.2576	4	sp|P00738;sp|P00739;	M[Oxidation]VSHHNLTTGATLINEQWLLTTAK	N4H5F0S0G0:4H
1169.817871	4	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N3H4F1S0G1:4H
918.4110107	4	sp|P01009;	YLGNATAIFFLPDEGK	N4H5F2S0G0:4H
1182.892212	3	sp|P00738;	VVLHPNYSQVDIGLIK	N4H4F0S1G0:3H
1061.18103	3	sp|P00738;	VVLHPNYSQVDIGLIK	N2H5F1S0G0:2H+ Na
1415.997803	3	sp|P00738;sp|P00739;	M[Oxidation]VSHHNLTTGATLINEQWLLTTAK	N3H4F2S0G0:3H
976.6748047	5	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N4H5F2S0G0:5H
1543.0811	3	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N5H4F0S0G0:3H

1434.6742	3	sp|P00738;sp|P00739;	MVSHHNLTTGATLINEQWLLTTAK	N4H5F0S0G0:3H
995.815979	6	sp|P01023;	TEVSSNHVLIYLDKVSNQTLSLFFTVLQDVPVR	N4H5F4S0G0:6H
1537.37439	3	sp|P00738;sp|P00739;	M[Oxidation]VSHHNLTTGATLINEQWLLTTAK	N4H5F2S0G0:3H
1179.82605	4	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N4H4F2S0G0:4H
1183.233032	3	sp|P00738;	VVLHPNYSQVDIGLIK	N4H4F2S0G0:3H
1531.7032	3	sp|P00738;sp|P00739;	MVSHHNLTTGATLINEQWLLTTAK	N4H4F1S0G1:3H
1198.583374	4	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N5H5F0S0G0:4H
1139.877319	3	sp|P00738;	VVLHPNYSQVDIGLIK	N4H5F0S0G0:3H
1085.864868	3	sp|P00738;	VVLHPNYSQVDIGLIK	N4H4F0S0G0:3H
1207.997192	2	sp|P01861;	EEQFNSTYR	N3H3F1S0G0:2H
1200.00415	2	sp|P01859;	EEQFNSTFR	N3H3F1S0G0:2H
1419.074829	2	sp|P01857;	EEQYNSTYR	N5H3F1S0G0:2H
1479.592163	2	sp|P01857;	EEQYNSTYR	N4H5F1S0G0:2H
1309.545166	2	sp|P01861;	EEQFNSTYR	N4H3F1S0G0:2H
1217.558716	3	sp|P01024;	TVLTPATNHM[Oxidation]GNVTFTIPANR	N2H6F0S0G0:3H
1200.005127	2	sp|P01859;	EEQFNSTFR	N3H3F1S0G0:2H
1244.509521	2	sp|P01857;	EEQYNSTYR	N4H3F0S0G0:2H
1471.5927	2	sp|P01861;	EEQFNSTYR	N4H5F1S0G0:2H
1217.222	3	sp|P01024;	TVLTPATNHM[Oxidation]GNVTFTIPANR	N2H6F0S0G0:3H
1217.226563	3	sp|P01024;	TVLTPATNHM[Oxidation]GNVTFTIPANR	N2H6F0S0G0:3H
949.6881714	4	sp|P01024;	TVLTPATNHMGNVTFTIPANR	N2H7F0S0G0:4H
953.6850586	4	sp|P01024;	TVLTPATNHM[Oxidation]GNVTFTIPANR	N2H7F0S0G0:4H
1202.060059	2	sp|P01877;	TPLTANITK	N4H3F1S0G0:2H
1344.578125	2	sp|P01871;	YKNNSDISSTR	N3H4F1S0G0:2H
1055.429932	3	sp|P02749;	LGNWSAMPSC[Carbamidomethyl]K	N4H4F1S0G1:3H
1020.453308	3	sp|P01877;	TPLTANITK	N5H4F1S1G0:3H
1055.505859	3	sp|P01024;	TVLTPATNHM[Oxidation]GNVTFTIPANR	N2H3F0S0G0:3H
1147.470703	3	sp|P04114;	YDFNSSMLYSTAK	N4H4F1S0G1:3H
1202.848633	3	sp|P0C0L4;sp|P0C0L5;	FSDGLESNSSTQFEVKK	N2H8F0S0G0:3H
1256.5257	3	sp|P0C0L4;sp|P0C0L5;	FSDGLESNSSTQFEVKK	N2H9F0S0G0:3H
1018.169617	3	sp|P00738;	VVLHPNYSQVDIGLIK	N3H4F0S0G0:3H
1074.80481	3	sp|P01877;	TPLTANITK	N5H5F3S0G0:3H
1115.20459	3	sp|P00738;	VVLHPNYSQVDIGLIK	N3H4F0S1G0:3H
1158.071899	4	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N5H4F0S0G0:4H
1180.077271	4	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N4H4F2S0G0:4H
855.1626587	4	sp|P00738;	VVLHPNYSQVDIGLIK	N4H5F0S0G0:4H
1020.482422	5	sp|P01009;	QLAHQSNSTNIFFSPVSIATAFAMLSLGTK	N4H5F2S0G0:5H
923.1567993	4	sp|P05546;	NLSMPLLPADFHK	N4H5F4S0G0:4H
913.15625	4	sp|P02790;	ALPQPQNVTSLLGC[Carbamidomethyl]TH	N4H4F1S0G1:4H
1086.194824	3	sp|P00738;	VVLHPNYSQVDIGLIK	N4H4F0S0G0:3H
1153.53479	4	sp|P00738;sp|P00739;	M[Oxidation]VSHHNLTTGATLINEQWLLTTAK	N4H5F2S0G0:4H
1538.037354	3	sp|P00738;sp|P00739;	M[Oxidation]VSHHNLTTGATLINEQWLLTTAK	N4H5F2S0G0:3H
1149.533813	4	sp|P00738;sp|P00739;	MVSHHNLTTGATLINEQWLLTTAK	N4H5F2S0G0:4H
1628.740234	3	sp|P00738;sp|P00739;	MVSHHNLTTGATLINEQWLLTTAK	N4H5F0S2G0:3H

1364.113647	2	sp|P01877;	TPLTANITK	N4H5F1S0G0:2H
1236.510742	2	sp|P01861;	EEQFNSTYR	N4H3F0S0G0:2H
1492.1073	2	sp|P01861;	EEQFNSTYR	N5H4F1S0G0:2H
1289.0247	2	sp|P01861;	EEQFNSTYR	N3H4F1S0G0:2H
1200.008301	2	sp|P01859;	EEQFNSTFR	N3H3F1S0G0:2H
1281.0249	2	sp|P01859;	EEQFNSTFR	N3H4F1S0G0:2H
1301.546	2	sp|P01859;	EEQFNSTFR	N4H3F1S0G0:2H
1228.516846	2	sp|P01859;	EEQFNSTFR	N4H3F0S0G0:2H
1325.531372	2	sp|P01857;	EEQYNSTYR	N4H4F0S0G0:2H
1411.0802	2	sp|P01861;	EEQFNSTYR	N5H3F1S0G0:2H
1317.5364	2	sp|P01857;	EEQYNSTYR	N4H3F1S0G0:2H
1463.592	2	sp|P01859;	EEQFNSTFR	N4H5F1S0G0:2H
1382.5657	2	sp|P01859;	EEQFNSTFR	N4H4F1S0G0:2H
1484.114	2	sp|P01859;	EEQFNSTFR	N5H4F1S0G0:2H
1317.53186	2	sp|P01857;	EEQYNSTYR	N4H3F1S0G0:2H
1611.1996	2	sp|P01877;	TPLTANITK	N5H5F1S1G0:2H
1217.224854	3	sp|P01024;	TVLTPATNHM[Oxidation]GNVTFTIPANR	N2H6F0S0G0:3H
1507.066528	3	sp|P05155;	VGQLQLSHNLSLVILVPQNLK	N4H5F2S1G0:3H
1271.240967	3	sp|P01024;	TVLTPATNHM[Oxidation]GNVTFTIPANR	N2H7F0S0G0:3H
1265.91272	3	sp|P01024;	TVLTPATNHMGNVTFTIPANR	N2H7F0S0G0:3H
1297.020996	2	sp|P01857;	EEQYNSTYR	N3H4F1S0G0:2H
1130.828003	3	sp|P02787;	C[Carbamidomethyl]GLVPVLAENYNK	N4H4F1S0G1:3H
1037.217163	4	sp|P01024;	TVLTPATNHM[Oxidation]GNVTFTIPANR	N3H6F0S1G0:4H
964.1516113	3	sp|P00738;	VVLHPNYSQVDIGLIK	N3H3F0S0G0:3H
1148.7695	4	sp|P04114;	QVLFLDTVYGNC[Carbamidomethyl]STHFTVK	N7H4F2S0G0:4H
1017.490112	4	sp|P00738;sp|P00739;	MVSHHNLTTGATLINEQWLLTTAK	N3H3F2S0G0:4H
1210.477051	3	sp|P08603;	MDGASNVTC[Carbamidomethyl]INSR	N4H5F0S2G0:3H
1229.873657	3	sp|P05546;	NLSMPLLPADFHK	N4H5F0S2G0:3H
1209.831665	4	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N3H6F0S1G0:4H
1825.3298	2	sp|P01024;	TVLTPATNHM[Oxidation]GNVTFTIPANR	N2H6F0S0G0:2H
1174.493286	3	sp|P10643;	INNDFNYEFYNSTWSYVK	N2H5F0S0G0:3H
1058.030151	4	sp|P05155;	VGQLQLSHNLSLVILVPQNLK	N4H5F2S0G0:4H
1021.483398	4	sp|P00738;sp|P00739;	M[Oxidation]VSHHNLTTGATLINEQWLLTTAK	N3H3F2S0G0:4H
1314.904175	3	sp|P02790;	ALPQPQNVTSLLGC[Carbamidomethyl]TH	N4H5F4S0G0:3H
1530.083252	3	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N4H5F0S0G0:3H
1230.845215	4	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N5H4F2S0G0:4H
1216.147705	3	sp|P08603;	M[Oxidation]DGASNVTC[Carbamidomethyl]INSR	N4H5F2S1G0:3H
1235.533203	3	sp|P05546;	NLSM[Oxidation]PLLPADFHK	N4H5F2S1G0:3H
1220.588501	4	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N4H5F2S0G0:4H
1115.202148	3	sp|P00738;	VVLHPNYSQVDIGLIK	N3H4F0S1G0:3H
1104.971802	4	sp|P01871;	GLTFQQNASSM[Oxidation]C[Carbamidomethyl]VPDQDT AIR	N4H5F3S0G0:4H
1634.0771	3	sp|P00738;sp|P00739;	M[Oxidation]VSHHNLTTGATLINEQWLLTTAK	N4H5F0S2G0:3H
1217.55127	3	sp|P01024;	TVLTPATNHM[Oxidation]GNVTFTIPANR	N2H6F0S0G0:3H
1134.973511	2	sp|P01861;	EEQFNSTYR	N3H3F0S0G0:2H

1297.024048	2	sp|P01857;	EEQYNSTYR	N3H4F1S0G0:2H
1403.0863	2	sp|P01859;	EEQFNSTFR	N5H3F1S0G0:2H
1398.564331	2	sp|P01857;	EEQYNSTYR	N4H4F1S0G0:2H
1500.1077	2	sp|P01857;	EEQYNSTYR	N5H4F1S0G0:2H
1301.5448	2	sp|P01859;	EEQFNSTFR	N4H3F1S0G0:2H
1398.569214	2	sp|P01857;	EEQYNSTYR	N4H4F1S0G0:2H
1659.6941	2	sp|P02790;	SWPAVGNC[Carbamidomethyl]SSALR	N4H4F1S0G1:2H
976.0667114	3	sp|P01859;	EEQFNSTFR	N4H5F1S0G0:3H
1509.6676	2	sp|P01877;	TPLTANITK	N4H4F2S0G1:2H
1024.415649	3	sp|P01861;	EEQFNSTYR	N4H4F1S1G0:3H
1097.442017	3	sp|P01857;	EEQYNSTYR	N5H4F1S1G0:3H
1073.104004	3	sp|P01859;	EEQFNSTFR	N4H4F2S0G1:3H
820.6875	3	sp|P04004;	NGSLFAFR	N3H4F0S1G0:3H
1228.19519	3	sp|P02787;	C[Carbamidomethyl]GLVPVLAENYNK	N4H5F2S1G0:3H
945.0957031	3	sp|P06681;	QSVPAHFVALNGSK	N2H6F0S0G0:3H
1321.241943	3	sp|P01009;	YLGNATAIFFLPDEGK	N4H5F2S1G0:3H
1100.47168	4	sp|P01871;	GLTFQQNASSMC[Carbamidomethyl]VPDQDTAIR	N4H4F2S0G1:4H
1224.211548	3	sp|P01009;	YLGNATAIFFLPDEGK	N4H5F2S0G0:3H
836.6552734	4	sp|P00738;	VVLHPNYSQVDIGLIK	N3H4F0S1G0:4H
1024.746948	3	sp|P05546;	DFVNASSK	N4H5F0S2G0:3H
1157.796265	3	sp|P02749;	LGNWSAM[Oxidation]PSC[Carbamidomethyl]K	N4H5F0S2G0:3H
1242.553345	3	sp|P01876;	LAGKPTHVNVSVVMAEVDGTC[Carbamidomethyl]Y	N2H6F0S0G0:3H
864.6004028	4	sp|P02749;	LGNWSAMPSC[Carbamidomethyl]K	N4H5F0S2G0:4H
1061.52002	3	sp|P00738;	VVLHPNYSQVDIGLIK	N3H3F2S0G0:3H
1023.76001	3	sp|P03952;	GVNFNVSK	N4H5F0S2G0:3H
1032.123657	3	sp|P27169;	HANWTLTPLK	N4H4F1S0G1:3H
1528.118	2	sp|P01859;	EEQFNSTFR	N4H4F1S1G0:2H
874.7070923	3	sp|P04004;	NGSLFAFR	N3H4F1S0G1:3H
986.7304077	3	sp|P01857;	EEQYNSTYR	N4H5F1S0G0:3H
1134.125366	3	sp|P43652;	DIENFNSTQK	N4H5F0S2G0:3H
1115.8042	3	sp|P01042;	LNAENNATFYFK	N4H4F1S0G1:3H
1147.810791	3	sp|P00450;	EHEGAIYPDNTTDFQR	N3H4F0S1G0:3H
1597.776978	3	sp|P01876;sp|P01877;	LSLHRPALEDLLLGSEANLTC[Carbamidomethyl]TLTGLR	N5H5F0S0G0:3H
1158.211426	3	sp|P01024;	TVLTPATNHMGNVTFTIPANR	N2H5F0S0G0:3H
1018.163696	3	sp|P00738;	VVLHPNYSQVDIGLIK	N3H4F0S0G0:3H
1311.5529	2	sp|P04004;	NGSLFAFR	N3H4F1S0G1:2H
1217.891357	3	sp|P01024;	TVLTPATNHM[Oxidation]GNVTFTIPANR	N2H6F0S0G0:3H
1212.55896	3	sp|P01024;	TVLTPATNHMGNVTFTIPANR	N2H6F0S0G0:3H
1695.732056	2	sp|P02787;	C[Carbamidomethyl]GLVPVLAENYNK	N4H4F1S0G1:2H
1557.6238	2	sp|P05155;	DTFVNASR	N4H5F0S2G0:2H
909.1743164	4	sp|P01024;	TVLTPATNHMGNVTFTIPANR	N2H6F0S0G0:4H

Table 3.9 Identified N-glycosylation sites in colon cancer and healthy human serum.

Charge	Spectrum mass	Protein ID	Sequence	Mod_Sites	Error (ppm)
2	2551.145	O95445	TELFSSSCPGGIMLJETGQGYQR	8,Carbamidomethyl[C](None);13,Oxidation[M](None);15,J_18O(None);	4.7896
2	2358.069	P01871	GLTFQQJASSMCVPDQDTAIR	7,J_18O(None);11,Oxidation[M](None);12,Carbamidomethyl[C](None);	4.2738
2	2360.246	P0DOX2	PALEDLLLGSEAJLTCTLTGLR	13,J_18O(None);16,Carbamidomethyl[C](None);	8.3267
2	2518.127	P02787	QQQHLFGSJVTDCSGJFCLFR	9,J_18O(None);13,Carbamidomethyl[C](None);18,Carbamidomethyl[C](None);	5.7606
2	2257.125	P05090	ADGTVNQIEGEATPVJLTEPAK	16,J_18O(None);	7.7275
2	2199.986	P02751	DQCIVDDITYNVJDTFHK	3,Carbamidomethyl[C](None);13,J_18O(None);	5.2441
2	1902.939	P02765	KVCQDCPLLAPLJDTR	3,Carbamidomethyl[C](None);6,Carbamidomethyl[C](None);13,J_18O(None);	5.0417
2	2317.194	P27169	VTQVYAEJGTVLQGSTVASVYK	8,J_18O(None);	5.5584
3	2640.254	P01042	HGIQYFNJNTQHSSLFMLNEVK	8,J_18O(None);17,Oxidation[M](None);	5.3726
3	2726.42	P05090	ADGTVNQIEGEATPVJLTEPAKLEVK	16,J_18O(None);	8.1602
3	2939.342	P12259	PYSIYPHGVTFSPYEDEVJSSFTSGR	19,J_18O(None);	5.6492
2	2235.155	P02751	LDAPTNLQFVJETDSTVLVR	11,J_18O(None);	7.3740
2	2368.158	P02765	AALAAFNAQNJGSNFQLEEISR	11,J_18O(None);	6.8669
3	3200.652	P01009	QLAHQSJSTNIFFSPVSIATAFAMLSLGTK	7,J_18O(None);24,Oxidation[M](None);	10.5429
2	2017.976	P04196	VIDFJCTTSSVSSALANTK	5,J_18O(None);6,Carbamidomethyl[C](None);	6.2746
2	2342.074	P01871	GLTFQQJASSMCVPDQDTAIR	7,J_18O(None);12,Carbamidomethyl[C](None);	4.3564
2	2812.365	Q14624	LPTQJITFQTESSVAEQEAEFQSPK	5,J_18O(None);	8.6696
3	3174.789	P01011	NVIFSPLSISTALAFLSLGAHJTTLTEILK	22,J_18O(None);	10.9979
2	2181.068	P01008	SLTFJETYQDISELVYGAK	5,J_18O(None);	9.0350
2	2400.236	P01023	GCVLLSYLJETVTVSASLESVR	2,Carbamidomethyl[C](None);9,J_18O(None);	5.9907
3	2097.011	P43251	NPVGLIGAEJATGETDPSHSK	10,J_18O(None);	6.2713
2	2258.164	P01024	TVLTPATNHMGJVTFTIPANR	12,J_18O(None);	6.7537
3	2614.249	P20851	LGHCPDPVLVNGEFSSSGPVJVSDK	4,Carbamidomethyl[C](None);21,J_18O(None);	5.9256
2	1885.908	P01008	LGACJDTLQQLMEVFK	4,Carbamidomethyl[C](None);5,J_18O(None);12,Oxidation[M](None);	8.6568
2	1951.78	P02751	HEEGHMLJCTCFGQGR	6,Oxidation[M](None);8,J_18O(None);9,Carbamidomethyl[C](None);11,Carbamidomethyl[C](None);	3.5860
2	2023.985	P08603	IPCSQPPQIEHGTIJSSR	3,Carbamidomethyl[C](None);15,J_18O(None);	4.7599
2	2025	P00450	ELHHLQEQJVSNAFLDK	9,J_18O(None);	3.9926
3	2826.419	Q96PD5	LEPVHLQLQCMSQEQLAQVAAJATK	10,Carbamidomethyl[C](None);11,Oxidation[M](None);22,J_18O(None);	6.2754
2	2235.154	P02751	LDAPTNLQFVJETDSTVLVR	11,J_18O(None);	6.9664
2	1974.079	P02748	AVJITSENLIDDVVSLIR	3,J_18O(None);	7.4359
3	3205.358	P02751	GGNSNGALCHFPFLYNNHJYTDCTSEGR	9,Carbamidomethyl[C](None);19,J_18O(None);23,Carbamidomethyl[C](None);	6.3678
3	2576.315	P36955	VTQJLTLIEESLTSEFIHDIDR	4,J_18O(None);	6.8839

2	2274.153	P01024	TVLTPATNHMGJVTFTIPANR	10,Oxidation[M](None);12,J_18O(None);	4.3889
2	1918.896	P02763	QDQCIYJTTYLNVQR	4,Carbamidomethyl[C](None);7,J_18O(None);	5.8528
2	2316.368	P05155	VGQLQLSHJLSLVILVPQNLK	9,J_18O(None);	6.3794
3	2360.251	P0DOX2	PALEDLLLGSEAJLTCTLTGLR	13,J_18O(None);16,Carbamidomethyl[C](None);	10.5671
2	2166.184	P01023	VSJQTLSLFFTVLQDVPVR	3,J_18O(None);	7.1439
2	2726.41	P05090	ADGTVNQIEGEATPVJLTEPAKLEVK	16,J_18O(None);	4.3097
2	1758.897	P01009	YLGJATAIFFLPDEGK	4,J_18O(None);	7.9772
2	1443.618	P08603	MDGASJVTCINSR	1,Oxidation[M](None);6,J_18O(None);9,Carbamidomethyl[C](None);	6.1013
3	2199.991	P02751	DQCIVDDITYNVJDTFHK	3,Carbamidomethyl[C](None);13,J_18O(None);	7.2278
2	1871.933	P12763	KLCPDCPLLAPLJDSR	3,Carbamidomethyl[C](None);6,Carbamidomethyl[C](None);13,J_18O(None);	5.1198
3	1886.013	P51884	LHINHNJLTESVGPLPK	7,J_18O(None);	5.9294
2	1612.735	P04275	MEACMLJGTVIGPGK	1,Oxidation[M](None);4,Carbamidomethyl[C](None);5,Oxidation[M](None);7,J_18O(None);	5.4454
3	3694.828	P01009	ADTHDEILEGLNFJLTEIPEAQIHEGFQELLR	14,J_18O(None);	4.3366
2	2384.102	P01871	STGKPTLYJVSLVMSDTAGTCY	9,J_18O(None);14,Oxidation[M](None);21,Carbamidomethyl[C](None);	5.8727
2	2384.106	P01871	STGKPTLYJVSLVMSDTAGTCY	9,J_18O(None);14,Oxidation[M](None);21,Carbamidomethyl[C](None);	7.5148
3	2589.131	P00751	SPYYJVSDEISFHCYDGYTLR	5,J_18O(None);14,Carbamidomethyl[C](None);	7.0352
3	2966.614	P0DOX2	LSLHRPALEDLLLGSEAJLTCTLTGLR	18,J_18O(None);21,Carbamidomethyl[C](None);	9.3966
2	1464.719	P00738	NLFLJHSEJATAK	5,J_18O(None);9,J_18O(None);	6.1193
3	3060.514	P00751	IVLDPSGSMNIYLVLDGSDSIGASJFTGAK	9,Oxidation[M](None);25,J_18O(None);	5.7471
3	2624.264	P01042	HGIQYFNJNTQHSSLFMLNEVK	8,J_18O(None);	7.5103
2	2128.136	P02763	CANLVPVPITJATLDQITGK	1,Carbamidomethyl[C](None);11,J_18O(None);	7.1429
3	3205.358	P02751	GGNSNGALCHFPFLYNNHJYTDCTSEGR	9,Carbamidomethyl[C](None);19,J_18O(None);23,Carbamidomethyl[C](None);	6.3678
3	2023.989	P08603	IPCSQPPQIEHGTIJSSR	3,Carbamidomethyl[C](None);15,J_18O(None);	7.0672
2	2366.139	P01876	LAGKPTHVJVSVVMAEVDGTCY	9,J_18O(None);14,Oxidation[M](None);21,Carbamidomethyl[C](None);	6.0089
2	1755.881	P01011	YTGJASALFILPDQDK	4,J_18O(None);	7.3319
3	3188.615	P00751	KIVLDPSGSMNIYLVLDGSDSIGASJFTGAK	10,Oxidation[M](None);26,J_18O(None);	7.3129
3	2966.583	P0DOX2	LSLHRPALEDLLLGSEAJLTCTLTGLR	18,J_18O(None);21,Carbamidomethyl[C](None);	-1.1552
3	3132.478	P02745	NPPMGGNVVIFDTVITNQEEPYQJHSGR	4,Oxidation[M](None);24,J_18O(None);	6.5440
3	3416.643	P01023	SLGNVJFTVSAEALESQELCGTEVPSVPEHGR	6,J_18O(None);20,Carbamidomethyl[C](None);	7.8328
3	2812.361	Q14624	LPTQJITFQTESSVAEQEAEFQSPK	5,J_18O(None);	7.1435
2	1869.905	P01008	LGACJDTLQQLMEVFK	4,Carbamidomethyl[C](None);5,J_18O(None);	4.3173
2	1777.811	P0C0L4	FSDGLESJSSTQFEVK	8,J_18O(None);	5.2773
2	1580.642	P08603	ISEEJETTCYMGK	5,J_18O(None);9,Carbamidomethyl[C](None);11,Oxidation[M](None);	5.2555
2	1686.833	P10909	LAJLTQGEDQYYLR	3,J_18O(None);	6.7813
3	3206.338	P02751	GGNSNGALCHFPFLYNNHJYTDCTSEGR	5,Deamidated[N](None);9,Carbamidomethyl[C](None);19,J_18O(None);23,Carbamidomethyl[C](None);	5.1005
2	1777.811	P0C0L4	FSDGLESJSSTQFEVK	8,J_18O(None);	5.7987

2	1400.797	P01833	VPGJVTAVLGETLK	4,J_18O(None);	6.7119
3	2551.151	O95445	TELFSSSCPGGIMLJETGQGYQR	8,Carbamidomethyl[C](None);13,Oxidation[M](None);15,J_18O(None);	7.1407
2	2368.11	P01871	STGKPTLYJVSLVMSDTAGTCY	9,J_18O(None);21,Carbamidomethyl[C](None);	7.2125
2	2366.14	P01876	LAGKPTHVJVSVVMAEVDGTCY	9,J_18O(None);14,Oxidation[M](None);21,Carbamidomethyl[C](None);	6.4557
2	1774.847	P02765	VCQDCPLLAPLJDTR	2,Carbamidomethyl[C](None);5,Carbamidomethyl[C](None);12,J_18O(None);	6.9138
2	2119.783	P05156	ACDGIJDCGDQSDELCCK	2,Carbamidomethyl[C](None);6,J_18O(None);8,Carbamidomethyl[C](None);16,Carbamidomethyl[C](None);17,Carbamidomethyl[C](None);	13.0825
3	3191.546	P19827	DKICDLLVANNHFAHFFAPQJLTNMNK	4,Carbamidomethyl[C](None);21,J_18O(None);25,Oxidation[M](None);	5.4206
2	1777.811	P0C0L4	FSDGLESJSSTQFEVK	8,J_18O(None);	5.7020
2	1923.864	P19652	QJQCFYJSSYLNVQR	4,Carbamidomethyl[C](None);7,J_18O(None);	5.1475
2	1439.798	P03952	IYSGILJLSDITK	7,J_18O(None);	7.2698
3	3674.783	P12763	RPTGEVYDIEIDTLETTCHVLDPTPLAJCSVR	18,Carbamidomethyl[C](None);28,J_18O(None);29,Carbamidomethyl[C](None);	7.3455
3	2274.157	P01024	TVLTPATNHMGJVTFTIPANR	10,Oxidation[M](None);12,J_18O(None);	6.1904
2	1443.618	P08603	MDGASJVTCINSR	1,Oxidation[M](None);6,J_18O(None);9,Carbamidomethyl[C](None);	6.0528
3	2698.397	P00738	MVSHHJLTTGATLINEQWLLTTAK	1,Oxidation[M](None);6,J_18O(None);	7.9317
4	2906.44	P00450	ELHHLQEQJVSNAFLDKGEFYIGSK	9,J_18O(None);	6.6951
3	2014.112	P51884	KLHINHNJLTESVGPLPK	8,J_18O(None);	7.4728
2	1758.9	P01009	YLGJATAIFFLPDEGK	4,J_18O(None);	9.3058
3	3019.58	P12763	VVHAVEVALATFNAESJGSYLQLVEISR	17,J_18O(None);	5.9402
3	2274.16	P01024	TVLTPATNHMGJVTFTIPANR	10,Oxidation[M](None);12,J_18O(None);	7.1842
3	2166.187	P01023	VSJQTLSLFFTVLQDVPVR	3,J_18O(None);	8.4259
3	1781.883	P01019	HLVIHJESTCEQLAK	6,J_18O(None);10,Carbamidomethyl[C](None);	5.1137
3	3206.341	P02751	GGNSNGALCHFPFLYNNHJYTDCTSEGR	5,Deamidated[N](None);9,Carbamidomethyl[C](None);19,J_18O(None);23,Carbamidomethyl[C](None);	6.2713
3	3112.712	P08185	AVLQLNEEGVDTAGSTGVTLJLTSKPIILR	21,J_18O(None);	4.3306
2	1743.843	P12763	LCPDCPLLAPLJDSR	2,Carbamidomethyl[C](None);5,Carbamidomethyl[C](None);12,J_18O(None);	8.3385
4	3839.69	P04217	EGDHEFLEVPEAQEDVEATFPVHQPGJYSCSYR	27,J_18O(None);30,Carbamidomethyl[C](None);	6.4026
4	3205.363	P02751	GGNSNGALCHFPFLYNNHJYTDCTSEGR	9,Carbamidomethyl[C](None);19,J_18O(None);23,Carbamidomethyl[C](None);	8.0450
2	3060.511	P00751	IVLDPSGSMNIYLVLDGSDSIGASJFTGAK	9,Oxidation[M](None);25,J_18O(None);	4.8482
4	2198.178	P51884	LGSFEGLVJLTFIHLQHNR	9,J_18O(None);	8.3301
2	1416.702	P01042	ITYSIVQTJCSK	9,J_18O(None);10,Carbamidomethyl[C](None);	6.4008
3	2948.43	P19827	ICDLLVANNHFAHFFAPQJLTNMNK	2,Carbamidomethyl[C](None);19,J_18O(None);23,Oxidation[M](None);	7.8228
4	3288.579	P02745	RNPPMGGNVVIFDTVITNQEEPYQJHSGR	5,Oxidation[M](None);25,J_18O(None);	6.2243
4	2906.441	P00450	ELHHLQEQJVSNAFLDKGEFYIGSK	9,J_18O(None);	6.9394
3	2063.986	Q02985	FVQGJSTEVACHPGYGLPK	5,J_18O(None);11,Carbamidomethyl[C](None);	5.7883
4	3205.361	P02751	GGNSNGALCHFPFLYNNHJYTDCTSEGR	9,Carbamidomethyl[C](None);19,J_18O(None);23,Carbamidomethyl[C](None);	7.2925
2	1987.01	P03952	IYPGVDFGGEELJVTFVK	13,J_18O(None);	7.6165

3	3417.636	P01023	SLGNVJFTVSAEALESQELCGTEVPSVPEHGR	6,J_18O(None);17,Deamidated[Q](None);20,Carbamidomethyl[C](None);	10.7015
4	3416.644	P01023	SLGNVJFTVSAEALESQELCGTEVPSVPEHGR	6,J_18O(None);20,Carbamidomethyl[C](None);	8.0883
2	2215.844	P07358	LLCNGDJDCGDQSDEANCR	3,Carbamidomethyl[C](None);7,J_18O(None);9,Carbamidomethyl[C](None);18,Carbamidomethyl[C](None);	12.3682
3	2274.164	P01024	TVLTPATNHMGJVTFTIPANR	10,Oxidation[M](None);12,J_18O(None);	9.2359
3	2682.397	P00738	MVSHHJLTTGATLINEQWLLTTAK	6,J_18O(None);	6.2664
3	3839.685	P04217	EGDHEFLEVPEAQEDVEATFPVHQPGJYSCSYR	27,J_18O(None);30,Carbamidomethyl[C](None);	5.1679
4	3132.478	P02745	NPPMGGNVVIFDTVITNQEEPYQJHSGR	4,Oxidation[M](None);24,J_18O(None);	6.5048
3	1733.912	P80108	LGTSLSSGHVLMJGTLK	12,Oxidation[M](None);13,J_18O(None);	7.7270
2	2384.101	P01871	STGKPTLYJVSLVMSDTAGTCY	9,J_18O(None);14,Oxidation[M](None);21,Carbamidomethyl[C](None);	5.5883
3	2199.991	P02751	DQCIVDDITYNVJDTFHK	3,Carbamidomethyl[C](None);13,J_18O(None);	7.4659
2	1941.881	P02790	SFDATTLDDJGTMLFFK	10,J_18O(None);13,Oxidation[M](None);	7.0566
4	2682.4	P00738	MVSHHJLTTGATLINEQWLLTTAK	6,J_18O(None);	7.1835
2	1743.845	P12763	LCPDCPLLAPLJDSR	2,Carbamidomethyl[C](None);5,Carbamidomethyl[C](None);12,J_18O(None);	9.3609
2	2248.042	P01011	TLJQSSDELQLSMGNAMFVK	3,J_18O(None);13,Oxidation[M](None);17,Oxidation[M](None);	2.7548
2	2138.973	P01011	HPNSPLDEEJLTQENQDR	10,J_18O(None);	12.3629
3	2025.006	P00450	ELHHLQEQJVSNAFLDK	9,J_18O(None);	6.7812
3	2358.076	P01871	GLTFQQJASSMCVPDQDTAIR	7,J_18O(None);11,Oxidation[M](None);12,Carbamidomethyl[C](None);	7.3026
3	2615.236	P20851	LGHCPDPVLVNGEFSSSGPVJVSDK	4,Carbamidomethyl[C](None);11,Deamidated[N](None);21,J_18O(None);	6.9653
2	2358.063	P01871	GLTFQQJASSMCVPDQDTAIR	7,J_18O(None);11,Oxidation[M](None);12,Carbamidomethyl[C](None);	1.7315
4	3674.78	P12763	RPTGEVYDIEIDTLETTCHVLDPTPLAJCSVR	18,Carbamidomethyl[C](None);28,J_18O(None);29,Carbamidomethyl[C](None);	6.6075
4	2966.608	P0DOX2	LSLHRPALEDLLLGSEAJLTCTLTGLR	18,J_18O(None);21,Carbamidomethyl[C](None);	7.4351
3	2316.375	P05155	VGQLQLSHJLSLVILVPQNLK	9,J_18O(None);	9.2800
2	1918.896	P02763	QDQCIYJTTYLNVQR	4,Carbamidomethyl[C](None);7,J_18O(None);	5.8695
2	1877.872	P01042	YNSQJQSNNQFVLYR	5,J_18O(None);	3.1877
2	1775.832	P02765	VCQDCPLLAPLJDTR	2,Carbamidomethyl[C](None);3,Deamidated[Q](None);5,Carbamidomethyl[C](None);12,J_18O(None);	7.2665
2	1545.677	P04114	YDFJSSMLYSTAK	4,J_18O(None);7,Oxidation[M](None);	7.2253
2	1783.025	P02763	LVPVPITJATLDQITGK	8,J_18O(None);	8.5652
2	1420.666	P29622	FLJDTMAVYEAK	3,J_18O(None);6,Oxidation[M](None);	7.9618
3	2128.138	P02763	CANLVPVPITJATLDQITGK	1,Carbamidomethyl[C](None);11,J_18O(None);	8.2720
4	3206.342	P02751	GGNSNGALCHFPFLYNNHJYTDCTSEGR	5,Deamidated[N](None);9,Carbamidomethyl[C](None);19,J_18O(None);23,Carbamidomethyl[C](None);	6.5570
4	2538.283	P01019	VYIHPFHLVIHJESTCEQLAK	12,J_18O(None);16,Carbamidomethyl[C](None);	5.3469
2	1427.863	P02750	LPPGLLAJFTLLR	8,J_18O(None);	8.3748
2	2068.952	P25311	DIVEYYJDSJGSHVLQGR	7,J_18O(None);	3.1620
2	1743.84	P12763	LCPDCPLLAPLJDSR	2,Carbamidomethyl[C](None);5,Carbamidomethyl[C](None);12,J_18O(None);	6.3159
4	3174.781	P01011	NVIFSPLSISTALAFLSLGAHJTTLTEILK	22,J_18O(None);	8.2758
3	2366.143	P01876	LAGKPTHVJVSVVMAEVDGTCY	9,J_18O(None);14,Oxidation[M](None);21,Carbamidomethyl[C](None);	7.6170

2	1925.752	P23142	DSSCGTGYELTEDJSCK	4,Carbamidomethyl[C](None);14,J_18O(None);16,Carbamidomethyl[C](None);	13.2696
3	3019.586	P12763	VVHAVEVALATFNAESJGSYLQLVEISR	17,J_18O(None);	7.9587
3	1791.752	P08185	MDPNAAYVJMSNHHR	1,Oxidation[M](None);9,J_18O(None);10,Oxidation[M](None);	5.3350
3	2810.424	Q96PD5	LEPVHLQLQCMSQEQLAQVAAJATK	10,Carbamidomethyl[C](None);22,J_18O(None);	6.5268
4	3191.552	P19827	DKICDLLVANNHFAHFFAPQJLTNMNK	4,Carbamidomethyl[C](None);21,J_18O(None);25,Oxidation[M](None);	7.2150
3	2274.156	P01024	TVLTPATNHMGJVTFTIPANR	10,Oxidation[M](None);12,J_18O(None);	5.5009
4	3206.348	P02751	GGNSNGALCHFPFLYNNHJYTDCTSEGR	9,Carbamidomethyl[C](None);16,Deamidated[N](None);19,J_18O(None);23,Carbamidomethyl[C](None);	8.2193
2	1783.026	P02763	LVPVPITJATLDQITGK	8,J_18O(None);	9.1698
5	2906.439	P00450	ELHHLQEQJVSNAFLDKGEFYIGSK	9,J_18O(None);	6.4429
2	1427.619	P08603	MDGASJVTCINSR	6,J_18O(None);9,Carbamidomethyl[C](None);	3.6943
2	2275.155	P01024	TVLTPATNHMGJVTFTIPANR	8,Deamidated[N](None);10,Oxidation[M](None);12,J_18O(None);	12.1895
3	3382.351	P02760	YFYJGTSMACETFQYGGCMGNGNNFVTEK	4,J_18O(None);8,Oxidation[M](None);10,Carbamidomethyl[C](None);18,Carbamidomethyl[C](None);19,Oxidation[M](None);	5.8394
2	1546.683	P05090	CIQAJYSLMENGK	1,Carbamidomethyl[C](None);5,J_18O(None);9,Oxidation[M](None);	4.4424
2	1374.746	P02763	PITJATLDQITGK	4,J_18O(None);	7.0726
2	2197.117	P35858	AGAFLGLTNVAVMJLSGNCLR	13,Oxidation[M](None);14,J_18O(None);19,Carbamidomethyl[C](None);	8.3651
3	2275.144	P01024	TVLTPATNHMGJVTFTIPANR	8,Deamidated[N](None);10,Oxidation[M](None);12,J_18O(None);	7.3296
2	1743.846	P12763	LCPDCPLLAPLJDSR	2,Carbamidomethyl[C](None);5,Carbamidomethyl[C](None);12,J_18O(None);	9.5318
2	1918.895	P02763	QDQCIYJTTYLNVQR	4,Carbamidomethyl[C](None);7,J_18O(None);	5.1472
2	1492.698	P19823	GAFISJFSMTVDGK	6,J_18O(None);9,Oxidation[M](None);	7.1059
3	2069.939	P25311	DIVEYYJDSJGSHVLQGR	7,J_18O(None);16,Deamidated[Q](None);	4.7199
3	1743.839	P12763	LCPDCPLLAPLJDSR	2,Carbamidomethyl[C](None);5,Carbamidomethyl[C](None);12,J_18O(None);	5.8325
3	2366.142	P01876	LAGKPTHVJVSVVMAEVDGTCY	9,J_18O(None);14,Oxidation[M](None);21,Carbamidomethyl[C](None);	7.0972
3	2518.131	P02787	QQQHLFGSJVTDCSGJFCLFR	9,J_18O(None);13,Carbamidomethyl[C](None);18,Carbamidomethyl[C](None);	7.3149
3	2198.179	P51884	LGSFEGLVJLTFIHLQHNR	9,J_18O(None);	8.3792
4	3995.796	P04217	REGDHEFLEVPEAQEDVEATFPVHQPGJYSCSYR	28,J_18O(None);31,Carbamidomethyl[C](None);	7.3890
3	2400.237	P01023	GCVLLSYLJETVTVSASLESVR	2,Carbamidomethyl[C](None);9,J_18O(None);	6.5619
2	1407.665	P02790	SWPAVGJCSSALR	7,J_18O(None);8,Carbamidomethyl[C](None);	5.3706
3	1369.704	P07996	VVJSTTGPGEHLR	3,J_18O(None);	6.2072
2	1918.899	P02763	QDQCIYJTTYLNVQR	4,Carbamidomethyl[C](None);7,J_18O(None);	7.2156
2	1592.731	Q03591	LQNNENJISCVER	7,J_18O(None);10,Carbamidomethyl[C](None);	5.5075
3	2366.142	P01876	LAGKPTHVJVSVVMAEVDGTCY	9,J_18O(None);14,Oxidation[M](None);21,Carbamidomethyl[C](None);	7.1302
3	1902.943	P02765	KVCQDCPLLAPLJDTR	3,Carbamidomethyl[C](None);6,Carbamidomethyl[C](None);13,J_18O(None);	7.0790
2	1929.857	P04275	GQVYLQCGTPCJLTCR	7,Carbamidomethyl[C](None);11,Carbamidomethyl[C](None);12,J_18O(None);15,Carbamidomethyl[C](None);	3.6635
3	1871.013	Q9UGM5	VLYLAAYJCTLRPVSK	8,J_18O(None);9,Carbamidomethyl[C](None);	8.0090
2	1477.664	P43652	YAEDKFJETTEK	7,J_18O(None);	4.1640

3	3019.587	P12763	VVHAVEVALATFNAESJGSYLQLVEISR	17,J_18O(None);	8.1461
3	2589.132	P00751	SPYYJVSDEISFHCYDGYTLR	5,J_18O(None);14,Carbamidomethyl[C](None);	7.5817
3	3416.64	P01023	SLGNVJFTVSAEALESQELCGTEVPSVPEHGR	6,J_18O(None);20,Carbamidomethyl[C](None);	6.9586
4	3207.327	P02751	GGNSNGALCHFPFLYNNHJYTDCTSEGR	5,Deamidated[N](None);9,Carbamidomethyl[C](None);16,Deamidated[N](None);19,J_18O(None);23,Carbamidomethyl[C](None);	6.7976
2	2018.962	P04196	VIDFJCTTSSVSSALANTK	5,J_18O(None);6,Carbamidomethyl[C](None);17,Deamidated[N](None);	7.1383
2	2094.129	Q96PD5	GFGVAIVGJYTAALPTEAALR	9,J_18O(None);	8.4503
4	3674.798	P12763	RPTGEVYDIEIDTLETTCHVLDPTPLAJCSVR	18,Carbamidomethyl[C](None);28,J_18O(None);29,Carbamidomethyl[C](None);	11.5070
3	1686.832	P10909	LAJLTQGEDQYYLR	3,J_18O(None);	5.9834
2	1434.687	P01042	LNAENJATFYFK	6,J_18O(None);	6.1003
3	2181.066	P01008	SLTFJETYQDISELVYGAK	5,J_18O(None);	8.0823
4	2640.258	P01042	HGIQYFNJNTQHSSLFMLNEVK	8,J_18O(None);17,Oxidation[M](None);	7.1391
3	1918.896	P02763	QDQCIYJTTYLNVQR	4,Carbamidomethyl[C](None);7,J_18O(None);	5.9138
2	1434.687	P01042	LNAENJATFYFK	6,J_18O(None);	6.0668
4	2949.408	P19827	ICDLLVANNHFAHFFAPQJLTNMNK	2,Carbamidomethyl[C](None);8,Deamidated[N](None);19,J_18O(None);23,Oxidation[M](None);	5.7849
3	3674.787	P12763	RPTGEVYDIEIDTLETTCHVLDPTPLAJCSVR	18,Carbamidomethyl[C](None);28,J_18O(None);29,Carbamidomethyl[C](None);	8.2824
3	2250.189	P03952	LQAPLJYTEFQKPICLPSK	6,J_18O(None);15,Carbamidomethyl[C](None);	7.4198
2	1321.654	O75882	IDSTGJVTNELR	6,J_18O(None);	4.8265
3	1555.956	P02750	KLPPGLLAJFTLLR	9,J_18O(None);	6.8472
2	1791.749	P08185	MDPNAAYVJMSNHHR	1,Oxidation[M](None);9,J_18O(None);10,Oxidation[M](None);	3.6596
2	3021.594	P12763	VVHAVEVALATFNAESJGSYLQLVEISR	13,Deamidated[N](None);17,J_18O(None);22,Deamidated[Q](None);	21.1627
2	1783.026	P02763	LVPVPITJATLDQITGK	8,J_18O(None);	9.4457
2	1738.88	P02790	ALPQPQJVTSLLGCTH	7,J_18O(None);14,Carbamidomethyl[C](None);	6.9090
2	1107.596	P0C0L4	GLJVTLSSTGR	3,J_18O(None);	6.0880
3	2258.166	P01024	TVLTPATNHMGJVTFTIPANR	12,J_18O(None);	7.7448
2	1642.789	P00450	AGLQAFFQVQECJK	12,Carbamidomethyl[C](None);13,J_18O(None);	6.9175
3	2641.237	P01042	HGIQYFNJNTQHSSLFMLNEVK	7,Deamidated[N](None);8,J_18O(None);17,Oxidation[M](None);	5.1495
3	2641.241	P01042	HGIQYFNJNTQHSSLFMLNEVK	8,J_18O(None);9,Deamidated[N](None);17,Oxidation[M](None);	6.4701
2	2966.601	P0DOX2	LSLHRPALEDLLLGSEAJLTCTLTGLR	18,J_18O(None);21,Carbamidomethyl[C](None);	4.8146
3	4804.257	P02751	ITTTPTNGQQGNSLEEVVHADQSSCTFDJLSPGLEYJVSVYTVK	12,Deamidated[N](None);25,Carbamidomethyl[C](None);37,J_18O(None);	5.6198
2	2236.153	P02751	LDAPTNLQFVJETDSTVLVR	6,Deamidated[N](None);11,J_18O(None);	13.6833
4	2714.272	P00734	YPHKPEIJSTTHPGADLQEJFCR	8,J_18O(None);22,Carbamidomethyl[C](None);	7.4981
5	3361.454	P02751	GGNSNGALCHFPFLYNNHJYTDCTSEGRR	9,Carbamidomethyl[C](None);19,J_18O(None);23,Carbamidomethyl[C](None);	4.7039
2	1777.811	P0C0L4	FSDGLESJSSTQFEVK	8,J_18O(None);	5.3082
2	2129.116	P02763	CANLVPVPITJATLDQITGK	1,Carbamidomethyl[C](None);3,Deamidated[N](None);11,J_18O(None);	5.3694
2	1935.86	P05155	NPJATSSSSQDPESLQDR	3,J_18O(None);	9.8261

2	2384.101	P01871	STGKPTLYJVSLVMSDTAGTCY	9,J_18O(None);14,Oxidation[M](None);21,Carbamidomethyl[C](None);	5.2905
2	1642.788	P00450	AGLQAFFQVQECJK	12,Carbamidomethyl[C](None);13,J_18O(None);	6.2357
5	3543.85	P01009	YLGJATAIFFLPDEGKLQHLENELTHDIITK	4,J_18O(None);	6.8200
2	1783.027	P02763	LVPVPITJATLDQITGK	8,J_18O(None);	9.5770
2	1287.616	P01871	YKJNSDISSTR	3,J_18O(None);	7.7577
2	1296.617	P55058	VSJVSCQASVSR	3,J_18O(None);6,Carbamidomethyl[C](None);	5.5321
3	1898.83	P07358	EYESYSDFERJVTEK	11,J_18O(None);	6.5056
3	2366.141	P01876	LAGKPTHVJVSVVMAEVDGTCY	9,J_18O(None);14,Oxidation[M](None);21,Carbamidomethyl[C](None);	6.6877
2	1580.636	P08603	ISEEJETTCYMGK	5,J_18O(None);9,Carbamidomethyl[C](None);11,Oxidation[M](None);	1.4823
3	2235.158	P02751	LDAPTNLQFVJETDSTVLVR	11,J_18O(None);	8.6437
3	1464.719	P00738	NLFLJHSEJATAK	5,J_18O(None);9,J_18O(None);	6.4565
2	1687.816	P10909	LAJLTQGEDQYYLR	3,J_18O(None);10,Deamidated[Q](None);	6.0901
2	1738.877	P02790	ALPQPQJVTSLLGCTH	7,J_18O(None);14,Carbamidomethyl[C](None);	5.3074
2	2250.189	P03952	LQAPLJYTEFQKPICLPSK	6,J_18O(None);15,Carbamidomethyl[C](None);	7.3789
2	1996.899	P01876	PTHVJVSVVMAEVDGTCY	5,J_18O(None);10,Oxidation[M](None);17,Carbamidomethyl[C](None);	5.5977
4	2025.004	P00450	ELHHLQEQJVSNAFLDK	9,J_18O(None);	5.8820
2	1738.883	P02790	ALPQPQJVTSLLGCTH	7,J_18O(None);14,Carbamidomethyl[C](None);	8.9943
4	3674.787	P12763	RPTGEVYDIEIDTLETTCHVLDPTPLAJCSVR	18,Carbamidomethyl[C](None);28,J_18O(None);29,Carbamidomethyl[C](None);	8.2995
3	2906.463	P00450	ELHHLQEQJVSNAFLDKGEFYIGSK	9,J_18O(None);	14.6535
3	2518.137	P02787	QQQHLFGSJVTDCSGJFCLFR	9,J_18O(None);13,Carbamidomethyl[C](None);18,Carbamidomethyl[C](None);	9.8910
3	1457.772	P06681	QSVPAHFVALJGSK	11,J_18O(None);	5.7183
3	1471.751	P02749	VYKPSAGJNSLYR	8,J_18O(None);	5.7224
2	1434.687	P01042	LNAENJATFYFK	6,J_18O(None);	5.9323
3	1895.843	P00450	EHEGAIYPDJTTDFQR	10,J_18O(None);	7.2991
2	1555.957	P02750	KLPPGLLAJFTLLR	9,J_18O(None);	7.5439
2	1314.597	P05156	FLNJGTCTAEGK	4,J_18O(None);7,Carbamidomethyl[C](None);	6.6013
4	2486.16	P05160	KEHETCLAPELYNGJYSTTQK	6,Carbamidomethyl[C](None);15,J_18O(None);	8.3752
3	3401.692	P01011	YTGJASALFILPDQDKMEEVEAMLLPETLK	4,J_18O(None);17,Oxidation[M](None);23,Oxidation[M](None);	8.6145
2	1886.91	P01008	LGACJDTLQQLMEVFK	4,Carbamidomethyl[C](None);5,J_18O(None);10,Deamidated[Q](None);12,Oxidation[M](None);	17.9866
3	2259.148	P01024	TVLTPATNHMGJVTFTIPANR	8,Deamidated[N](None);12,J_18O(None);	7.0186
3	2384.106	P01871	STGKPTLYJVSLVMSDTAGTCY	9,J_18O(None);14,Oxidation[M](None);21,Carbamidomethyl[C](None);	7.3306
3	2366.14	P01876	LAGKPTHVJVSVVMAEVDGTCY	9,J_18O(None);14,Oxidation[M](None);21,Carbamidomethyl[C](None);	6.1535
3	1871.939	P12763	KLCPDCPLLAPLJDSR	3,Carbamidomethyl[C](None);6,Carbamidomethyl[C](None);13,J_18O(None);	7.8192
4	2966.614	P0DOX2	LSLHRPALEDLLLGSEAJLTCTLTGLR	18,J_18O(None);21,Carbamidomethyl[C](None);	9.1512
3	2714.267	P00734	YPHKPEIJSTTHPGADLQEJFCR	8,J_18O(None);22,Carbamidomethyl[C](None);	5.9633
3	1923.867	P19652	QJQCFYJSSYLNVQR	4,Carbamidomethyl[C](None);7,J_18O(None);	6.6527
2	2384.102	P01871	STGKPTLYJVSLVMSDTAGTCY	9,J_18O(None);14,Oxidation[M](None);21,Carbamidomethyl[C](None);	5.8320
3	1783.023	P02763	LVPVPITJATLDQITGK	8,J_18O(None);	7.7559

4	2014.109	P51884	KLHINHNJLTESVGPLPK	8,J_18O(None);	6.0538
3	1895.838	P00450	EHEGAIYPDJTTDFQR	10,J_18O(None);	4.6549
2	1841.94	Q9UK55	LPYQGJATMLVVLMEK	6,J_18O(None);9,Oxidation[M](None);14,Oxidation[M](None);	6.9101
3	1609.811	P04114	FVEGSHJSTVSLTTK	7,J_18O(None);	9.5962
2	1916.931	P04114	FEVDSPVYJATWSASLK	9,J_18O(None);	7.6805
5	4903.479	P02750	SDHGSSISCQPPAEIPGYLPADTVHLAVEFFJLTHLPANLLQGASK	9,Carbamidomethyl[C](None);32,J_18O(None);39,Deamidated[N](None);	13.4123
3	1673.815	P55058	GKEGHFYYJISEVK	9,J_18O(None);	5.6894
2	2975.33	O75882	MPSQAPTGNFYPQPLLJSSMCLEDSR	1,Oxidation[M](None);17,J_18O(None);20,Oxidation[M](None);21,Carbamidomethyl[C](None);	6.7266
3	1758.895	P01009	YLGJATAIFFLPDEGK	4,J_18O(None);	6.8583
2	2425.995	P07996	VSCPIMPCSJATVPDGECCPR	3,Carbamidomethyl[C](None);6,Oxidation[M](None);8,Carbamidomethyl[C](None);10,J_18O(None);18,Carbamidomethyl[C](None);19,Carbamidomethyl[C](None);	6.3586
3	2023.988	P08603	IPCSQPPQIEHGTIJSSR	3,Carbamidomethyl[C](None);15,J_18O(None);	6.3083
4	1951.783	P02751	HEEGHMLJCTCFGQGR	6,Oxidation[M](None);8,J_18O(None);9,Carbamidomethyl[C](None);11,Carbamidomethyl[C](None);	5.3833
3	3416.639	P01023	SLGNVJFTVSAEALESQELCGTEVPSVPEHGR	6,J_18O(None);20,Carbamidomethyl[C](None);	6.6931
2	2384.102	P01871	STGKPTLYJVSLVMSDTAGTCY	9,J_18O(None);14,Oxidation[M](None);21,Carbamidomethyl[C](None);	6.0195
3	2368.157	P02765	AALAAFNAQNJGSNFQLEEISR	11,J_18O(None);	6.5439
3	2384.103	P01871	STGKPTLYJVSLVMSDTAGTCY	9,J_18O(None);14,Oxidation[M](None);21,Carbamidomethyl[C](None);	6.4225
2	1434.688	P01042	LNAENJATFYFK	6,J_18O(None);	6.2836
2	961.5504	P0DOX2	TPLTAJITK	6,J_18O(None);	5.6336
2	1434.689	P01042	LNAENJATFYFK	6,J_18O(None);	7.2064
3	1477.666	P43652	YAEDKFJETTEK	7,J_18O(None);	5.6427
2	1434.687	P01042	LNAENJATFYFK	6,J_18O(None);	6.1003
2	1743.841	P12763	LCPDCPLLAPLJDSR	2,Carbamidomethyl[C](None);5,Carbamidomethyl[C](None);12,J_18O(None);	6.6858
2	2275.141	P01024	TVLTPATNHMGJVTFTIPANR	8,Deamidated[N](None);10,Oxidation[M](None);12,J_18O(None);	6.1020
2	1136.568	P17936	GLCVJASAVSR	3,Carbamidomethyl[C](None);5,J_18O(None);	6.1061
2	1277.654	Q08380	AAIPSALDTJSSK	10,J_18O(None);	5.9672
2	1434.686	P01042	LNAENJATFYFK	6,J_18O(None);	4.9425
3	1721.847	Q96IY4	QVHFFVJASDVDNVK	7,J_18O(None);	5.0748
3	2342.079	P01871	GLTFQQJASSMCVPDQDTAIR	7,J_18O(None);12,Carbamidomethyl[C](None);	6.5839
3	1555.958	P02750	KLPPGLLAJFTLLR	9,J_18O(None);	7.7444
3	1461.73	P00738	NLFLJHSEJATAK	5,J_18O(None);	5.8725
3	1774.846	P02765	VCQDCPLLAPLJDTR	2,Carbamidomethyl[C](None);5,Carbamidomethyl[C](None);12,J_18O(None);	6.5223
3	2366.14	P01876	LAGKPTHVJVSVVMAEVDGTCY	9,J_18O(None);14,Oxidation[M](None);21,Carbamidomethyl[C](None);	6.1581
2	1774.847	P02765	VCQDCPLLAPLJDTR	2,Carbamidomethyl[C](None);5,Carbamidomethyl[C](None);12,J_18O(None);	7.0981
3	2486.15	P05160	KEHETCLAPELYNGJYSTTQK	6,Carbamidomethyl[C](None);15,J_18O(None);	4.3569
2	1374.747	P02763	PITJATLDQITGK	4,J_18O(None);	8.0786
4	3175.558	P19827	DKICDLLVANNHFAHFFAPQJLTNMNK	4,Carbamidomethyl[C](None);21,J_18O(None);	7.6314
2	2698.399	P00738	MVSHHJLTTGATLINEQWLLTTAK	1,Oxidation[M](None);6,J_18O(None);	8.6937

3	2225.112	P03951	LETTVJYTDSQRPICLPSK	6,J_18O(None);15,Carbamidomethyl[C](None);	5.1566
2	1198.554	P43652	DIENFJSTQK	6,J_18O(None);	5.8904
4	2682.4	P00738	MVSHHJLTTGATLINEQWLLTTAK	6,J_18O(None);	7.1835
3	1642.816	P01859	TKPREEQFJSTFR	9,J_18O(None);	5.8071
3	2949.41	P19827	ICDLLVANNHFAHFFAPQJLTNMNK	2,Carbamidomethyl[C](None);9,Deamidated[N](None);19,J_18O(None);23,Oxidation[M](None);	6.5539
2	1454.752	P03952	LQAPLJYTEFQK	6,J_18O(None);	7.2390
3	2347.185	P29622	SQILEGLGFJLTELSESDVHR	10,J_18O(None);	7.9444
5	3191.547	P19827	DKICDLLVANNHFAHFFAPQJLTNMNK	4,Carbamidomethyl[C](None);21,J_18O(None);25,Oxidation[M](None);	5.9351
2	2368.157	P02765	AALAAFNAQNJGSNFQLEEISR	11,J_18O(None);	6.7327
3	2949.43	P19827	ICDLLVANNHFAHFFAPQJLTNMNK	2,Carbamidomethyl[C](None);8,Deamidated[N](None);19,J_18O(None);23,Oxidation[M](None);	13.4226
4	3674.788	P12763	RPTGEVYDIEIDTLETTCHVLDPTPLAJCSVR	18,Carbamidomethyl[C](None);28,J_18O(None);29,Carbamidomethyl[C](None);	8.7118
4	3753.901	P01023	IITILEEEMJVSVCGLYTYGKPVPGHVTVSICR	9,Oxidation[M](None);10,J_18O(None);14,Carbamidomethyl[C](None);32,Carbamidomethyl[C](None);	5.9173
4	2949.434	P19827	ICDLLVANNHFAHFFAPQJLTNMNK	2,Carbamidomethyl[C](None);8,Deamidated[N](None);19,J_18O(None);23,Oxidation[M](None);	14.5336
4	2356.092	P02751	DQCIVDDITYNVJDTFHKR	3,Carbamidomethyl[C](None);13,J_18O(None);	6.9318
4	2949.413	P19827	ICDLLVANNHFAHFFAPQJLTNMNK	2,Carbamidomethyl[C](None);8,Deamidated[N](None);19,J_18O(None);23,Oxidation[M](None);	7.5890
3	1885.906	P01008	LGACJDTLQQLMEVFK	4,Carbamidomethyl[C](None);5,J_18O(None);12,Oxidation[M](None);	7.7093
4	2949.413	P19827	ICDLLVANNHFAHFFAPQJLTNMNK	2,Carbamidomethyl[C](None);8,Deamidated[N](None);19,J_18O(None);23,Oxidation[M](None);	7.6727
3	2366.137	P01876	LAGKPTHVJVSVVMAEVDGTCY	9,J_18O(None);14,Oxidation[M](None);21,Carbamidomethyl[C](None);	4.8569
2	1783.023	P02763	LVPVPITJATLDQITGK	8,J_18O(None);	7.6045
2	1192.507	P01857	EEQYJSTYR	5,J_18O(None);	5.3861
3	3674.76	P12763	RPTGEVYDIEIDTLETTCHVLDPTPLAJCSVR	18,Carbamidomethyl[C](None);28,J_18O(None);29,Carbamidomethyl[C](None);	1.0847
3	2518.118	P02787	QQQHLFGSJVTDCSGJFCLFR	9,J_18O(None);13,Carbamidomethyl[C](None);18,Carbamidomethyl[C](None);	2.2283
4	3674.78	P12763	RPTGEVYDIEIDTLETTCHVLDPTPLAJCSVR	18,Carbamidomethyl[C](None);28,J_18O(None);29,Carbamidomethyl[C](None);	6.5786
2	1192.507	P01857	EEQYJSTYR	5,J_18O(None);	5.5589
2	1547.668	P05090	CIQAJYSLMENGK	1,Carbamidomethyl[C](None);5,J_18O(None);9,Oxidation[M](None);11,Deamidated[N](None);	5.3539
2	1570.869	P02763	PVPITJATLDQITGK	6,J_18O(None);	7.4207
3	3674.775	P12763	RPTGEVYDIEIDTLETTCHVLDPTPLAJCSVR	18,Carbamidomethyl[C](None);28,J_18O(None);29,Carbamidomethyl[C](None);	5.1230
4	1886.011	P51884	LHINHNJLTESVGPLPK	7,J_18O(None);	5.2534
3	2274.155	P01024	TVLTPATNHMGJVTFTIPANR	10,Oxidation[M](None);12,J_18O(None);	5.1342
3	1798.01	P00738	VVLHPJYSQVDIGLIK	6,J_18O(None);	5.8559
2	1411.815	P19652	LVPVPITJATLDR	8,J_18O(None);	7.7333
3	2235.156	P02751	LDAPTNLQFVJETDSTVLVR	11,J_18O(None);	7.8187
4	1781.885	P01019	HLVIHJESTCEQLAK	6,J_18O(None);10,Carbamidomethyl[C](None);	6.4825
3	1871.934	P12763	KLCPDCPLLAPLJDSR	3,Carbamidomethyl[C](None);6,Carbamidomethyl[C](None);13,J_18O(None);	5.5606
4	3674.783	P12763	RPTGEVYDIEIDTLETTCHVLDPTPLAJCSVR	18,Carbamidomethyl[C](None);28,J_18O(None);29,Carbamidomethyl[C](None);	7.2927
4	3019.585	P12763	VVHAVEVALATFNAESJGSYLQLVEISR	17,J_18O(None);	7.6987
2	961.5516	P0DOX2	TPLTAJITK	6,J_18O(None);	6.9159

2	1434.688	P01042	LNAENJATFYFK	6,J_18O(None);	6.3665
2	1160.517	P01859	EEQFJSTFR	5,J_18O(None);	5.7922
2	1672.88	Q9NZP8	PVTPIAQJQTTLGSSR	8,J_18O(None);	3.1700
3	2275.145	P01024	TVLTPATNHMGJVTFTI PANR	10,Oxidation[M](None);12,J_18O(None);20,Deamidated[N](None);	7.9485
3	2192.076	Q02985	KFVQGJSTEVACHPGY GLPK	6,J_18O(None);12,Carbamidomethyl[C](None);	3.2590
3	2779.353	P20160	FVJVTVTPEDQCRPNJV CTGVLTR	3,J_18O(None);12,Carbamidomethyl[C](None);18,Carba midomethyl[C](None);	5.0735
4	2091.895	P02751	RHEEGHMLJCTCFGQG R	9,J_18O(None);10,Carbamidomethyl[C](None);12,Carba midomethyl[C](None);	7.3866
3	2366.142	P01876	LAGKPTHVJVSVVMAE VDGTCY	9,J_18O(None);14,Oxidation[M](None);21,Carbamidome thyl[C](None);	7.0883
2	1192.507	P01857	EEQYJSTYR	5,J_18O(None);	5.7677
3	2366.141	P01876	LAGKPTHVJVSVVMAE VDGTCY	9,J_18O(None);14,Oxidation[M](None);21,Carbamidome thyl[C](None);	6.9024
4	4673.273	Q14624	NQALJLSLAYSFVTPLT SMVVTKPDDQEQSQVA EKPMEGESR	5,J_18O(None);19,Oxidation[M](None);37,Oxidation[M] (None);	5.2967
3	2518.131	P02787	QQQHLFGSJVTDCSGJF CLFR	9,J_18O(None);13,Carbamidomethyl[C](None);18,Carba midomethyl[C](None);	7.2530
2	1160.517	P01859	EEQFJSTFR	5,J_18O(None);	5.3011
2	1176.513	P01860	EEQYJSTFR	5,J_18O(None);	6.5643
2	1160.517	P01859	EEQFJSTFR	5,J_18O(None);	5.4700
3	2366.138	P01876	LAGKPTHVJVSVVMAE VDGTCY	9,J_18O(None);14,Oxidation[M](None);21,Carbamidome thyl[C](None);	5.5039
2	1738.873	P02790	ALPQPQJVTSLLGCTH	7,J_18O(None);14,Carbamidomethyl[C](None);	3.2717
2	1160.518	P01859	EEQFJSTFR	5,J_18O(None);	6.2929
4	2948.435	P19827	ICDLLVANNHFAHFFAP QJLTNMNK	2,Carbamidomethyl[C](None);19,J_18O(None);23,Oxidat ion[M](None);	9.4803
3	2350.144	P01876	LAGKPTHVJVSVVMAE VDGTCY	9,J_18O(None);21,Carbamidomethyl[C](None);	5.9473
2	1176.512	P01861	EEQFJSTYR	5,J_18O(None);	5.7186
2	1738.879	P02790	ALPQPQJVTSLLGCTH	7,J_18O(None);14,Carbamidomethyl[C](None);	6.2236
3	1743.842	P12763	LCPDCPLLAPLJDSR	2,Carbamidomethyl[C](None);5,Carbamidomethyl[C](No ne);12,J_18O(None);	7.6848
3	3020.586	P12763	VVHAVEVALATFNAES JGSYLQLVEISR	17,J_18O(None);22,Deamidated[Q](None);	13.300 7
2	1192.506	P01857	EEQYJSTYR	5,J_18O(None);	4.9274
2	1743.842	P12763	LCPDCPLLAPLJDSR	2,Carbamidomethyl[C](None);5,Carbamidomethyl[C](No ne);12,J_18O(None);	7.4365
3	3518.678	P12763	PTGEVYDIEIDTLETTC HVLDPTPLAJCSVR	17,Carbamidomethyl[C](None);27,J_18O(None);28,Carb amidomethyl[C](None);	6.4754
2	1777.837	P0C0L4	FSDGLESJSSTQFEVK	8,J_18O(None);	20.260 0
2	2128.108	P02763	CANLVPVPITJATLDQIT GK	1,Carbamidomethyl[C](None);11,J_18O(None);	- 5.8446
2	1154.511	O75882	CIJQSICEK	1,Carbamidomethyl[C](None);3,J_18O(None);7,Carbami domethyl[C](None);	3.6327
2	1071.544	P20851	KTLFCJASK	5,Carbamidomethyl[C](None);6,J_18O(None);	5.0609
3	1738.88	P02790	ALPQPQJVTSLLGCTH	7,J_18O(None);14,Carbamidomethyl[C](None);	6.9965
3	2518.129	P02787	QQQHLFGSJVTDCSGJF CLFR	9,J_18O(None);13,Carbamidomethyl[C](None);18,Carba midomethyl[C](None);	6.6958
4	2715.251	P00734	YPHKPEIJSTTHPGADL QEJFCR	8,J_18O(None);18,Deamidated[Q](None);22,Carbamido methyl[C](None);	5.7899
2	1176.513	P01861	EEQFJSTYR	5,J_18O(None);	6.1572
5	3362.453	P02751	GGNSNGALCHFPFLYN NHJYTDCTSEGRR	5,Deamidated[N](None);9,Carbamidomethyl[C](None);1 9,J_18O(None);23,Carbamidomethyl[C](None);	9.1722
4	2641.244	P01042	HGIQYFNJNTQHSSLFM LNEVK	8,J_18O(None);9,Deamidated[N](None);17,Oxidation[M] (None);	7.8849

4	3192.527	P19827	DKICDLLVANNHFAHF FAPQJLTNMNK	4,Carbamidomethyl[C](None);11,Deamidated[N](None); 21,J_18O(None);25,Oxidation[M](None);	4.5819
2	1055.53	P03951	GINYJSSVAK	5,J_18O(None);	4.7701
2	2547.031	P02790	CSDGWSFDATTLDDJG TMLFFK	1,Carbamidomethyl[C](None);15,J_18O(None);18,Oxidat ion[M](None);	######
2	1968.105	P02763	ANLVPVPITJATLDQITG K	10,J_18O(None);	7.7196
3	1738.881	P02790	ALPQPQJVTSLLGCTH	7,J_18O(None);14,Carbamidomethyl[C](None);	7.4887
3	2317.197	P27169	VTQVYAEJGTVLQGST VASVYK	8,J_18O(None);	7.2363
3	2274.156	P01024	TVLTPATNHMGJVTFTI PANR	10,Oxidation[M](None);12,J_18O(None);	5.4996
2	1529.684	P04114	YDFJSSMLYSTAK	4,J_18O(None);	8.4704
3	2366.14	P01876	LAGKPTHVJVSVVMAE VDGTCY	9,J_18O(None);14,Oxidation[M](None);21,Carbamidome thyl[C](None);	6.2629
3	2723.316	P04114	ELCTISHIFIPAMGJITYD FSFK	3,Carbamidomethyl[C](None);13,Oxidation[M](None);15 ,J_18O(None);	6.6357
2	1743.842	P12763	LCPDCPLLAPLJDSR	2,Carbamidomethyl[C](None);5,Carbamidomethyl[C](No ne);12,J_18O(None);	7.2690
3	1342.698	Q96KN2	LVPHMJVSAVEK	5,Oxidation[M](None);6,J_18O(None);	4.7881
2	1642.813	P01859	TKPREEQFJSTFR	9,J_18O(None);	3.6511
2	1674.804	P01857	TKPREEQYJSTYR	9,J_18O(None);	4.4704
2	2384.101	P01871	STGKPTLYJVSLVMSDT AGTCY	9,J_18O(None);14,Oxidation[M](None);21,Carbamidome thyl[C](None);	5.3794
3	2371.946	P07358	LLCNGDJDCGDQSDEA NCRR	3,Carbamidomethyl[C](None);7,J_18O(None);9,Carbami domethyl[C](None);18,Carbamidomethyl[C](None);	11.629 3
2	1172.545	O75882	DLDMFIJASK	4,Oxidation[M](None);7,J_18O(None);	5.2424
2	1192.506	P01857	EEQYJSTYR	5,J_18O(None);	4.7002
2	2235.157	P02751	LDAPTNLQFVJETDSTV LVR	11,J_18O(None);	8.2549
5	3362.453	P02751	GGNSNGALCHFPFLYN NHJYTDCTSEGRR	9,Carbamidomethyl[C](None);16,Deamidated[N](None); 19,J_18O(None);23,Carbamidomethyl[C](None);	9.2602
3	1871.936	P12763	KLCPDCPLLAPLJDSR	3,Carbamidomethyl[C](None);6,Carbamidomethyl[C](No ne);13,J_18O(None);	6.4639
3	2195.061	P02748	FSYSKJETYQLFLSYSS K	6,J_18O(None);	8.2818
3	2366.14	P01876	LAGKPTHVJVSVVMAE VDGTCY	9,J_18O(None);14,Oxidation[M](None);21,Carbamidome thyl[C](None);	6.3242
4	3674.788	P12763	RPTGEVYDIEIDTLETT CHVLDPTPLAJCSVR	18,Carbamidomethyl[C](None);28,J_18O(None);29,Carb amidomethyl[C](None);	8.6228
4	1717.794	P02768	QEPERJECFLQHK	6,J_18O(None);8,Carbamidomethyl[C](None);	5.5833
2	2197.116	P35858	AGAFLGLTNVAVMJLS GNCLR	13,Oxidation[M](None);14,J_18O(None);19,Carbamidom ethyl[C](None);	7.7843
2	1212.589	P06681	TMFPJLTDVR	2,Oxidation[M](None);5,J_18O(None);	6.5727
3	2519.113	P02787	QQQHLFGSJVTDCSGJF CLFR	3,Deamidated[Q](None);9,J_18O(None);13,Carbamidom ethyl[C](None);18,Carbamidomethyl[C](None);	6.4662
2	1112.59	P07339	GSLSYLJVTR	7,J_18O(None);	6.4651
4	3152.424	P04275	TEPMQVALHCTJGSVV YHEVLNAMECK	4,Oxidation[M](None);10,Carbamidomethyl[C](None);12 ,J_18O(None);24,Oxidation[M](None);26,Carbamidomet hyl[C](None);	6.4500
4	3674.781	P12763	RPTGEVYDIEIDTLETT CHVLDPTPLAJCSVR	18,Carbamidomethyl[C](None);28,J_18O(None);29,Carb amidomethyl[C](None);	6.8608
2	1783.022	P02763	LVPVPITJATLDQITGK	8,J_18O(None);	6.6937
3	2366.143	P01876	LAGKPTHVJVSVVMAE VDGTCY	9,J_18O(None);14,Oxidation[M](None);21,Carbamidome thyl[C](None);	7.4687
2	988.4931	Q9UK55	ETFFJLSK	5,J_18O(None);	6.0193
3	1488.697	P55058	EGHFYYJISEVK	7,J_18O(None);	5.5176
2	1445.666	P01857	PREEQYJSTYR	7,J_18O(None);	8.2149
3	2518.13	P02787	QQQHLFGSJVTDCSGJF CLFR	9,J_18O(None);13,Carbamidomethyl[C](None);18,Carba midomethyl[C](None);	7.0401

3	1896.825	P00450	EHEGAIYPDJTTDFQR	10,J_18O(None);15,Deamidated[Q](None);	6.4044
2	2166.184	P01023	VSJQTLSLFFTVLQDVP VR	3,J_18O(None);	7.1439
2	1686.833	P10909	LAJLTQGEDQYYLR	3,J_18O(None);	6.7813
2	2236.157	P02751	LDAPTNLQFVJETDSTV LVR	8,Deamidated[Q](None);11,J_18O(None);	15.087 5
2	1161.502	P01859	EEQFJSTFR	3,Deamidated[Q](None);5,J_18O(None);	6.3383
2	950.4359	P10909	EDALJETR	5,J_18O(None);	5.0808
3	1738.875	P02790	ALPQPQJVTSLLGCTH	7,J_18O(None);14,Carbamidomethyl[C](None);	4.3551
2	961.5504	P0DOX2	TPLTAJITK	6,J_18O(None);	5.6752
3	2518.133	P02787	QQQHLFGSJVTDCSGJF CLFR	9,J_18O(None);13,Carbamidomethyl[C](None);18,Carba midomethyl[C](None);	8.2529
2	1192.506	P01857	EEQYJSTYR	5,J_18O(None);	4.9232
4	3765.03	P01023	TEVSSNHVLIYLDKVSJ QTLSLFFTVLQDVPVR	17,J_18O(None);	8.2414
2	1555.958	P02750	KLPPGLLAJFTLLR	9,J_18O(None);	8.1140
3	1287.614	P01871	YKJNSDISSTR	3,J_18O(None);	6.1377
2	1738.878	P02790	ALPQPQJVTSLLGCTH	7,J_18O(None);14,Carbamidomethyl[C](None);	5.8428
5	2957.402	P00734	SRYPHKPEIJSTTHPGA DLQEJFCR	10,J_18O(None);24,Carbamidomethyl[C](None);	6.0543
2	1434.689	P01042	LNAENJATFYFK	6,J_18O(None);	7.0287
5	2107.884	P02751	RHEEGHMLJCTCFGQG R	7,Oxidation[M](None);9,J_18O(None);10,Carbamidomet hyl[C](None);12,Carbamidomethyl[C](None);	4.9443
2	1160.517	P01859	EEQFJSTFR	5,J_18O(None);	5.6285
2	870.4137	P05546	DFVJASSK	4,J_18O(None);	5.5066
4	2274.159	P01024	TVLTPATNHMGJVTFTI PANR	10,Oxidation[M](None);12,J_18O(None);	6.8865
3	2603.255	P06315	ETTLTQSPAFMSATPGD KVJISCK	6,Deamidated[Q](None);11,Oxidation[M](None);20,J_18 O(None);23,Carbamidomethyl[C](None);	17.396 3
3	1877.876	P01042	YNSQJQSNNQFVLYR	5,J_18O(None);	5.4849
2	2017.976	P04196	VIDFJCTTSSVSSALANT K	5,J_18O(None);6,Carbamidomethyl[C](None);	6.3415
3	1413.675	P01859	PREEQFJSTFR	7,J_18O(None);	7.7635
3	2366.146	P01876	LAGKPTHVJVSVVMAE VDGTCY	9,J_18O(None);14,Oxidation[M](None);21,Carbamidome thyl[C](None);	8.6833
2	1010.451	O75882	GICJSSDVR	3,Carbamidomethyl[C](None);4,J_18O(None);	5.5916
2	1592.731	Q03591	LQNNENJISCVER	7,J_18O(None);10,Carbamidomethyl[C](None);	5.5075
4	3674.782	P12763	RPTGEVYDIEIDTLETT CHVLDPTPLAJCSVR	18,Carbamidomethyl[C](None);28,J_18O(None);29,Carb amidomethyl[C](None);	7.0279
2	1177.495	P01861	EEQFJSTYR	3,Deamidated[Q](None);5,J_18O(None);	5.1210
3	1445.66	P01857	PREEQYJSTYR	7,J_18O(None);	4.3862
3	2274.158	P01024	TVLTPATNHMGJVTFTI PANR	10,Oxidation[M](None);12,J_18O(None);	6.5681
3	1674.803	P01857	TKPREEQYJSTYR	9,J_18O(None);	3.4786
3	2535.152	O95445	TELFSSSCPGGIMLJETG QGYQR	8,Carbamidomethyl[C](None);15,J_18O(None);	5.6139
3	1895.84	P00450	EHEGAIYPDJTTDFQR	10,J_18O(None);	5.6587
2	1413.702	O95497	LTGVAGJYTVCQK	7,J_18O(None);11,Carbamidomethyl[C](None);	6.4200
2	2384.101	P01871	STGKPTLYJVSLVMSDT AGTCY	9,J_18O(None);14,Oxidation[M](None);21,Carbamidome thyl[C](None);	5.2531
3	2366.141	P01876	LAGKPTHVJVSVVMAE VDGTCY	9,J_18O(None);14,Oxidation[M](None);21,Carbamidome thyl[C](None);	6.5757
3	2200.989	P02751	DQCIVDDITYNVJDTFH K	3,Carbamidomethyl[C](None);11,Deamidated[N](None); 13,J_18O(None);	13.503 9
4	3694.826	P01009	ADTHDEILEGLNFJLTEI PEAQIHEGFQELLR	14,J_18O(None);	3.6944

3	2204.2	Q8WZ75	IQLEJVTLLNPDPAEGPKPR	5,J_18O(None);	8.7075
2	1160.517	P01859	EEQFJSTFR	5,J_18O(None);	5.9370
4	4803.285	P02751	ITTTPTNGQQGNSLEEVVHADQSSCTFDJLSPGLEYJVSVYTVK	25,Carbamidomethyl[C](None);37,J_18O(None);	8.0151
2	1192.509	P01857	EEQYJSTYR	5,J_18O(None);	7.3257
4	2933.411	P19827	ICDLLVANNHFAHFFAPQJLTNMNK	2,Carbamidomethyl[C](None);8,Deamidated[N](None);19,J_18O(None);	5.1960
2	1192.506	P01857	EEQYJSTYR	5,J_18O(None);	4.9098
2	1176.51	P01860	EEQYJSTFR	5,J_18O(None);	4.3502
3	2366.14	P01876	LAGKPTHVJVSVVMAEVDGTCY	9,J_18O(None);14,Oxidation[M](None);21,Carbamidomethyl[C](None);	6.2203
3	2350.15	P01876	LAGKPTHVJVSVVMAEVDGTCY	9,J_18O(None);21,Carbamidomethyl[C](None);	8.3429
3	1658.811	P01861	TKPREEQFJSTYR	9,J_18O(None);	5.7354
3	1784.006	P02763	LVPVPITJATLDQITGK	8,J_18O(None);13,Deamidated[Q](None);	7.1121
5	2714.266	P00734	YPHKPEIJSTTHPGADLQEJFCR	8,J_18O(None);22,Carbamidomethyl[C](None);	5.3823
3	3293.607	P08185	VTISGVYDLGDVLEEMGIADLFTNQAJFSR	16,Oxidation[M](None);27,J_18O(None);	9.3372
4	2966.593	P0DOX2	LSLHRPALEDLLLGSEAJLTCTLTGLR	18,J_18O(None);21,Carbamidomethyl[C](None);	2.1826
3	2384.13	P01871	STGKPTLYJVSLVMSDTAGTCY	9,J_18O(None);14,Oxidation[M](None);21,Carbamidomethyl[C](None);	17.6433
2	1528.751	P07711	YSVAJDTGFVDIPK	5,J_18O(None);	6.2005
3	1755.881	P01011	YTGJASALFILPDQDK	4,J_18O(None);	7.0944
2	825.4652	P13671	VLJFTTK	3,J_18O(None);	5.9966
4	4634.209	P04217	FQSPAGTEALFELHJISVADSAJYSCVYVDLKPPFGGSAPSER	15,J_18O(None);23,J_18O(None);26,Carbamidomethyl[C](None);	6.2229
2	912.4369	P05155	DTFVJASR	5,J_18O(None);	6.8147
3	3184.623	P01009	QLAHQSJSTNIFFSPVSIATAFAMLSLGTK	7,J_18O(None);	-0.0279
3	2199.991	P02751	DQCIVDDITYNVJDTFHK	3,Carbamidomethyl[C](None);13,J_18O(None);	7.5891
2	1192.507	P01857	EEQYJSTYR	5,J_18O(None);	5.3912
2	1471.75	P02749	VYKPSAGJNSLYR	8,J_18O(None);	4.9825
4	1598.729	P49221	JLTVDTYVNENGEK	1,J_18O(None);	-9.0028
3	2366.146	P01876	LAGKPTHVJVSVVMAEVDGTCY	9,J_18O(None);14,Oxidation[M](None);21,Carbamidomethyl[C](None);	8.7970
3	1609.805	P04114	FVEGSHJSTVSLTTK	7,J_18O(None);	6.0318
5	3996.796	P04217	REGDHEFLEVPEAQEDVEATFPVHQPGJYSCSYR	14,Deamidated[Q](None);28,J_18O(None);31,Carbamidomethyl[C](None);	11.5650
2	1775.826	P02765	VCQDCPLLAPLJDTR	2,Carbamidomethyl[C](None);3,Deamidated[Q](None);5,Carbamidomethyl[C](None);12,J_18O(None);	4.2442
4	3674.787	P12763	RPTGEVYDIEIDTLETTCHVLDPTPLAJCSVR	18,Carbamidomethyl[C](None);28,J_18O(None);29,Carbamidomethyl[C](None);	8.4587
3	2395.237	Q14624	KAFITJFSMIIDGMTYPGIIK	6,J_18O(None);9,Oxidation[M](None);14,Oxidation[M](None);	8.0439
3	3674.775	P12763	RPTGEVYDIEIDTLETTCHVLDPTPLAJCSVR	18,Carbamidomethyl[C](None);28,J_18O(None);29,Carbamidomethyl[C](None);	5.2093
5	3674.787	P12763	RPTGEVYDIEIDTLETTCHVLDPTPLAJCSVR	18,Carbamidomethyl[C](None);28,J_18O(None);29,Carbamidomethyl[C](None);	8.5289
4	2698.397	P00738	MVSHHJLTTGATLINEQWLLTTAK	1,Oxidation[M](None);6,J_18O(None);	8.0136
2	867.4503	P03952	GVNFJVSK	5,J_18O(None);	5.4435

4 CHAPTER 4. DEPOLYMERIZATION AND CHARACTERIZATION OF STREPTOCOCCUS PNEUMONIAE TYPE 8 POLYSACCHARIDE

4.1 Abstract

Although there have been decades of research on streptococcus pneumoniae, it is still among the leading cause of infectious disease in the world. As a type of capsular polysaccharide (CPS) of streptococcus pneumoniae, pneumococcal polysaccharides are essential components for colonization and virulence in mammalian hosts. This study aimed to characterize the CPS structure of type 8 streptococcus pneumoniae, which is one of the most fatal serotypes. In this work, heparinase I&III was used to successfully digest pneumococcal type 8 polysaccharide (Pn8P). We characterized the oligosaccharide generated from the enzymatic depolymerization of Pn8P by size exclusion chromatography, mass spectrometry and nuclear magnetic resonance. This is the first study to enzymatically depolymerize and characterize Pn8P.

4.2 Introduction

Pneumococcal disease is a globally spread disease which can cause fever, coughing, joint pain, chills and ear pain. The streptococcus pneumoniae is usually alive in the nose and throat of 5–10% of healthy adults and 20–40% of healthy children.[1] Pneumococcal polysaccharide is a carbohydrate polymer that surrounds the surface of streptococcus pneumoniae, which has been widely used as vaccines against of pneumococcal disease.[2-4] The polysaccharide chain is necessary for pneumococcal colonization and virulence in biological organisms, which makes it an ideal immunoprophylactic antigen.[5, 6] These polysaccharide chains prevent streptococcus pneumoniae from complement-mediated clearance from the lung.[7] Streptococcus pneumoniae expresses over 90 immunological polysaccharides structures adhere to the surface of cell, each with an unique repeating unite.[8] As with most of the other pneumococcal polysaccharides, the repeat unit structure

of pneumococcal type 8 polysaccharide (Pn8P) was deduced from genetic analysis to be -4-GlcA-β1,4-Glc-β1,4-Glc-α1,4-Gal-α1-, but there has been no study to characterize its structure directly.[9, 10]

Previously, size exclusion chromatography (SEC) has been reported to characterize some capsular polysaccharides which have been used in vaccines given their stability and immunogenicity.[11, 12] Three types of pneumococcal polysaccharides and their conjugates haven been well studied by size exclusion chromatography in different pH and ionic strength conditions.[13] Because UV detection is limited in carbohydrate analysis due to the lack of chromophores, the evaporative light scattering detector (ELSD) and refractive index detector (RID) are generally used in the detection of glycans. Both methods have been reported to perform well on polysaccharide analysis.[14, 15] Although mass spectrometry (MS) has been widely used in glycomics and glycoproteomics to study oligosaccharides including *N*-glycan, *O*-glycan and others, long carbohydrate polymer chains can not be efficiently ionized in MS without major fragmentation taking place. However, recently short polysaccharides such as low molecular weight heparins (LMWHs) have been fully sequenced by high resolution MS and tandem mass techniques.[16, 17]

Normally, to determine the structure of polysaccharides by MS, depolymerization is generally used to degrade long carbohydrate polymer chains into low molecular weight oligosaccharides.[18] Enzymatic depolymerization of certain polysaccharides using glycosaminoglycan (GAG) degrading enzymes were proven to be applicable.[19] Depending on the enzymatic mechanism, glycosaminoglycan enzymes can be sorted into two categories: hydrolases and lyases.[20] Lyases cleave the linkage between hexosamine and glucuronic acid via a β elimination reaction to produce a 4,5-double bond between C-4 and C-5 on the glucuronic acid

residue.[21] Hydrolase cleave the hexuronic acid-hexosamine bond through hydrolytic cleavage, while the glycosylic bond is hydrolyzed by the addition of a water molecule. (Figure 4.1) Chondroitinase ABC and heparinase are representative lyases, and hyaluronidase is widely used hydrolase.[22] Given high degrees of heterogeneity and high molecular weights, studies related to pneumococcal polysaccharide characterization is limited. For example, Middleton and coworkers depolymerized pneumococcal type 3 polysaccharide using a hydrolase called paenibacillus species 32352.[23] Herein, we present a novel enzymatic method to depolymerize Pn8P. HPLC-RI-MS and nuclear magnetic resonance (NMR) spectroscopy were used to corroborate the depolymerization and fully characterize the purified Pn8P repeating units.

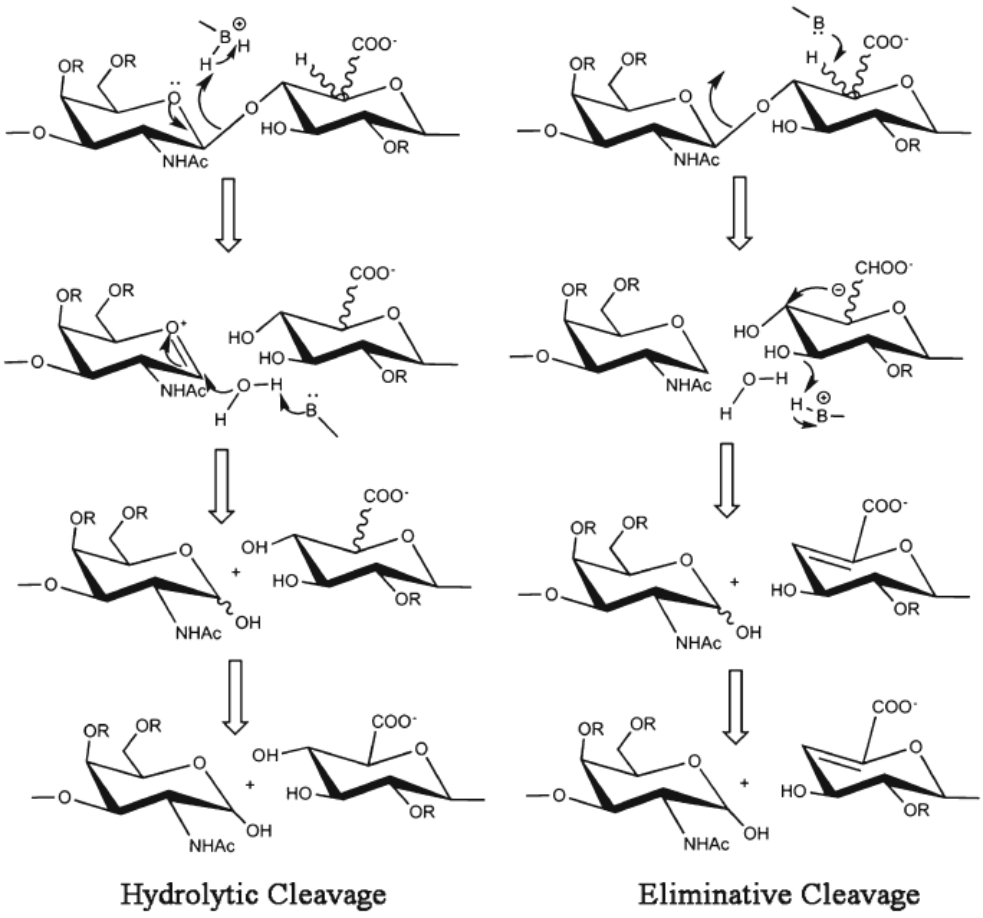


Figure 4.1 Mechanism of polysaccharide hydrolase and lyase.

4.3 Experimental

4.3.1 Materials

Pneumococcal type 8 polysaccharide (Pn8P) (American Type Culture Collection, Merck Sharp & Dohme deposit), chondroitinase ABC (Sigma Aldrich), hyaluronidase from bovine testes (Sigma Aldrich), heparinase I & III from Flavobacterium heparinum (Sigma Aldrich), chondroitin sulfate from shark cartilage (Sigma Aldrich), 1M Tris-HCl buffer pH 7.5 (Fisher Scientific) , sodium acetate (Sigma Aldrich), sodium chloride (Fisher Scientific), ammonium formate (ACROS Organics).

4.3.2 Enzymatic degradation of Pn8P

3mg of Chondroitin Sulfate was dissolve in 1ml of 50 mM Tris-HCl buffer, pH 7.0. A volume of 30µL of Chondroitin Sulfate 40 mU Chondroitinase ABC in 50 mM Tris-HCl buffer, 60 mM sodium acetate, at pH 8.0 and incubated for 48 hours at 37°C. A volume of 30µL of Chondroitin Sulfate was incubated with 40 U Hyaluronidase in 0.1 M sodium acetate buffer, 150 mM NaCl, at pH 5.0 and incubated for 48 hours at 37°C. 3mg of Chondroitin Sulfate was dissolve in 1ml of 0.1 M sodium acetate in DI water, at pH 6.5. A volume of 30Ml of heparin was incubated with Heparinase I&III in 0.1 M sodium acetate in DI water, at pH 6.5 for 48 hours at 37°C.

Pneumococcal type 8 polysaccharide (Pn8P) was dissolved to a concentration of 3 mg/mL, and 30µL of polysaccharide was incubated with 40 mU Chondroitinase ABC in 50 mM Tris-HCl buffer, 60 mM sodium acetate in DI water, at pH 8.0 for 48 hours at 37°C. A volume of 30µL of Pn8P was incubated with 40 U Hyaluronidase in buffer 0.1 M sodium acetate in DI water, 150 mM NaCl, at pH 5.0 for 48 hours at 37°C.Another 30µL of Pn8P was incubated with Heparinase I&III in 0.1 M sodium acetate in DI water, at pH 6.5 for 48 hours at 37°C. All enzymes were

inactivated by heat at 95°C for 5 min. The activities of all enzymes were tested by standard substrate.

4.3.3 Fast scan of depolymerization by HPLC-ELSD

The HPLC system used for the fast scan of the enzymatic reaction was Shimadzu Prominence UFLC coupled with a Sedere ELSD detector. The reaction solution was separated by using a Size Exclusion Column (Waters BEH SEC, 125Å, 1.7 μm, 4.6 mm × 150 mm) with a Size Exclusion guard column (Waters BEH SEC 125Å, 1.7 μm, 4.6 mm × 30 mm) using isocratic elution of 50 mM ammonium formate (pH 6.3) at a flow rate of 0.2 mL/min. Total run time was 20 minutes. Column oven and ELSD detector temperatures were kept at 60°C. Data was acquired and was processed using LabSolutions software (Shimadzu Corporation, USA).

4.3.4 Reaction condition optimization

The heparinase digestion was then optimized under different buffer conditions and different time range. First, 7.5 μL Pn8P was incubated with 0 U, 0.1 U, 0.25 U, 0.5 U, 0.75 U, 1 U heparinase I&III at 25°C for 30min, 1h, 4h, 8h, 24h, 30h, 48h, 55h in 20mM Tris-HCl buffer pH 7.0 with 100mM NaCl. Four buffer conditions were then monitored separately, which were 100 mM sodium acetate in DI water, pH 6.5, 20 mM Tris-HCl pH 6.5, 7.0, 7.5 with 100mM NaCl. The reaction was then tested at three different temperature 25°C, 37°C, 50°C. Finally, the reaction activators were tested at CaCl2, 10 mM CaCl2, 1 mM MgCl2, 10 mM MgCl2 conditions separately. The enzymes in all reaction solutions were inactivated by heated at 95°C for 5 min. The reaction was monitored by using a nanodrop 2000c spectrophotometer (Thermo Scientific) to measure the UV absorption at wavelength 232 nm, which would increase with the double bond formation on the non-reducing end of glucuronic acid.

4.3.5 HPLC-RI-MSn analysis of oligosaccharide

The Shimadzu Prominence UFLC coupled with a Refractive Index detector (Shimadzu, RID-20A) was used for analysis and separation of the oligosaccharides in the reaction solution. Size exclusion chromatography was performed on a SEC column (GE Health Care Superdex 200, 300 mm × 10 mm, 8.6 μm) by using isocratic elution of 70mM ammonium formate (pH 6.3) at a flow rate of 0.3 mL/min. Total run time was 80 minutes. Column and detector temperature were maintained at 30°C. Data was acquired and processed using LabSolutions software (Shimadzu Corporation, USA). As the oligosaccharide eluted, they were collected and lyophilized for further analysis. Oligosaccharides were dissolved in 0.1% formic acid and injected into an ESI-MS (LTQ-Orbitrap Elite mass spectrometer, Thermo Fisher). The ESI-MS was operated in both of positive and negative ion mode using the following settings: ion spray voltage 1.5 kv, ion source heater temperature 200°C. Precursor ion spectra was acquired in the Fourier Transform Mass Spectrometry (FTMS) analyzer with a resolution of 60000. Collision-induced dissociation (CID) fragment ions spectra were acquired in the ion trap analyzer with a resolution of 15,000. CID spectra were collected using an isolation width of 1 Da, collision energy of 35, an activation Q of 0.25, and an activation time of 10 ms. Target precursor ions were selected by multiple reaction monitor (MRM) mode for MS2, MS3 and MS4 fragmentation to further confirm the structure of the oligosaccharide. Data was acquired and processed using Xcalibur software.

4.4 Results and discussion

4.4.1 Discovery of target enzyme

Three different enzymes of GAG, including Chondroitinase ABC, heparinase and hyaluronidase were used to digest Pn8P. The three enzymatic reactions were monitored by the HPLC-ELSD fast scan method. Through a relatively short SEC column, both polysaccharide and

low molecular weight oligosaccharide eluted within 20 minutes. From the ELSD results, all of the Pn8P, chondroitin sulfate and heparin were eluted at 6.2 min, and the salts in the buffer system were eluted from 11min to 15min.

In the standard enzyme activity test, chondroitinase ABC and hyaluronidase completely digested chondroitin sulfate in 24 hours, and heparinase I&III depolymerized heparin in 24 hours. (Figure 4.7) Neither chondroitinase ABC (lyase) nor hyaluronidase (hydrolase) digested Pn8P, the target polysaccharide, while heparinase (lyase) did successfully digest Pn8P into low molecular weight oligosaccharides. (Figure 4.2)

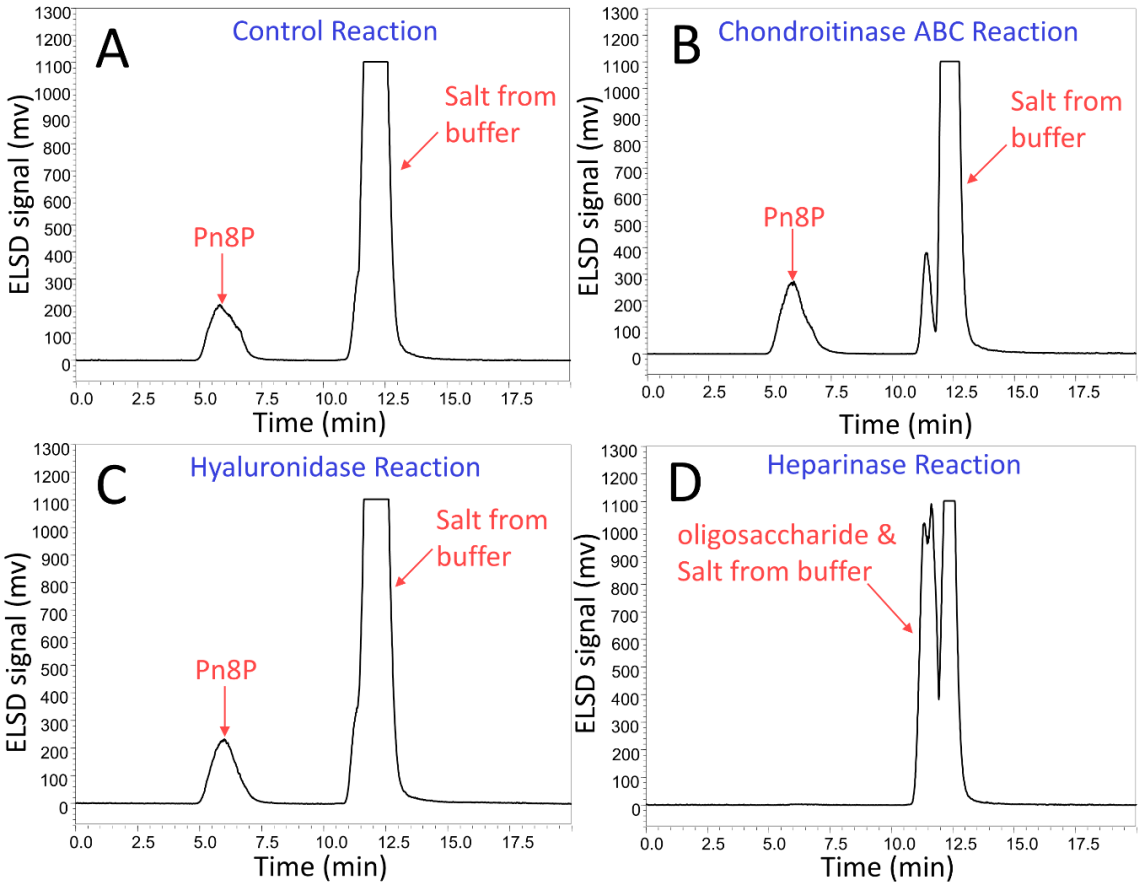


Figure 4.2 SEC-ELSD results of enzymatic reactions with Pn8P A. Control reaction; B. Chondroitinase ABC reaction with Pn8P; C. Hyaluronidase reaction with Pn8P; D. Heparinase reaction with Pn8P.

Heparinase is known to cleave the beta glycosidic bond between galactose and glucuronic acid.[22] The repeat unit structure of Pn8P was -4-GlcA-α1,4-Glc-β1,4-Glc-β1,4-Gal-α1-, theoretically it should be digested to the tetrasaccharide resulting in an unsaturated C-4 and C-5 bond on glucuronic acid. (Figure 4.3) After digestion, the structure of the polysaccharide repeat unit could be separated and further confirmed. Surprisingly, even though the linkage between galactose and glucuronic acid is α1,4-linkage,[10] the enzyme can still recognize the substrate and cleave the alpha glycosidic bond.

4.4.2 Enzyme activity optimization

According to the β-elimination mechanism of heparinase, a double bond is formed between C-4 and C-5 on the glucuronic acid residue. Therefore, the reaction solutions' UV absorption at 232 nm would increase as the digestion of Pn8P progresses and this could be used to measure the reactions progress. All of the UV 232 nm absorbance results from nanodrop were listed in supplemental materials.(Table 4.2-4.5) A "one factor at a time"(OFAT) set of experiments was conducted to determine the optimum conditions for digestion. First, after incubation with 0U, 0.1U, 0.25U, 0.5U, 0.75U, 1U heparinase I&III, Pneumococcal type 8 polysaccharide (Pn8P) could be completely degraded by at least 0.5U enzyme. 0.5 U enzyme was applied to all the rest optimizations. Secondly, different buffer systems were optimized at 25°C for 30min, 1h, 4h, 8h, 24h, 30h, 48h, 55h, the results showed decreasing value of degradation from pH 6.5~7.5 in Tris-HCl condition, and Tris-HCl pH 6.5 with 100mM NaCl had better degradation than 100mM sodium acetate. The buffer 20mM Tris-HCl pH 6.5 with 100mM NaCl was then applied to the later optimizations.

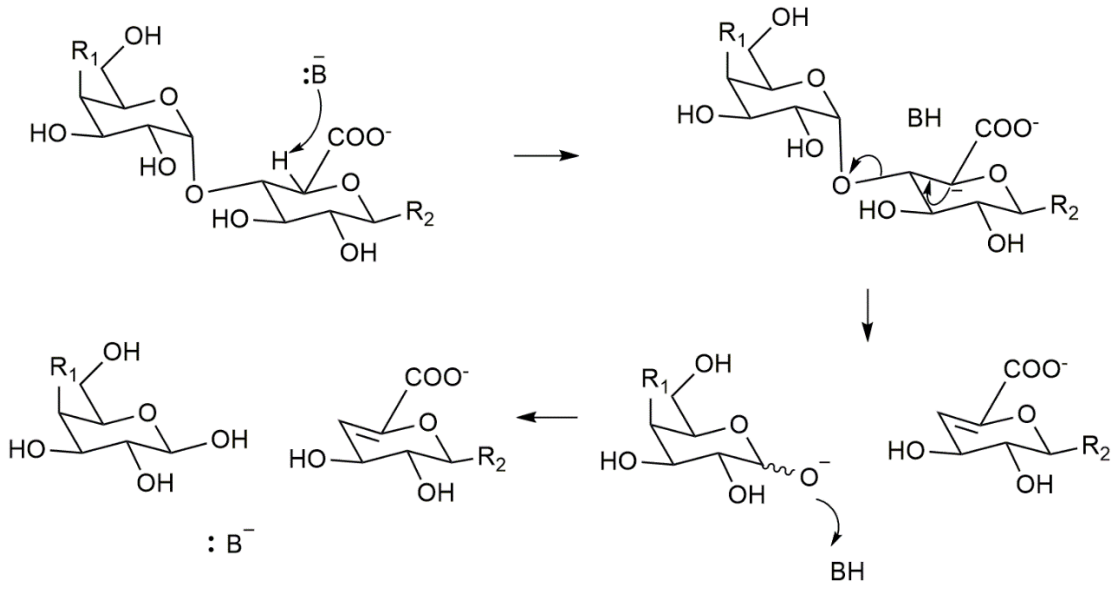


Figure 4.3 Beta eliminative mechanism of Pn8P digestion by heparinase.

In the reaction temperature optimization, the enzyme had better activity at 50°C than at 37°C from 30min to 30h. However, it stopped reacting from 30h at 50°C, while the reaction in 37°C generated more degraded oligosaccharides after 48h than 50°C's reaction. The temperature 37°C and buffer 20mM Tris-HCl pH 6.5 with 100mM NaCl were then applied to activators test. Different activators were added to promote the reaction, and the addition of 1mM Mg2+ promoted the best results. In the activators optimization, 1mM Ca2+ did not affect the reaction, 10mM Ca2+ had negative influence on the reaction, 10mM Mg2+ slightly improved the reaction, and 1mM Mg2+ significantly promoted the reaction. Enzyme concentration, reaction temperature and different buffer conditions were monitored. The optimized reaction condition was 45 μg Pn8P per unit enzyme, in 20 mM Tris-HCl pH 6.5, 1 mM Mg2+, buffer, incubation at 37°C for 48 hours. (Figure 4.4)

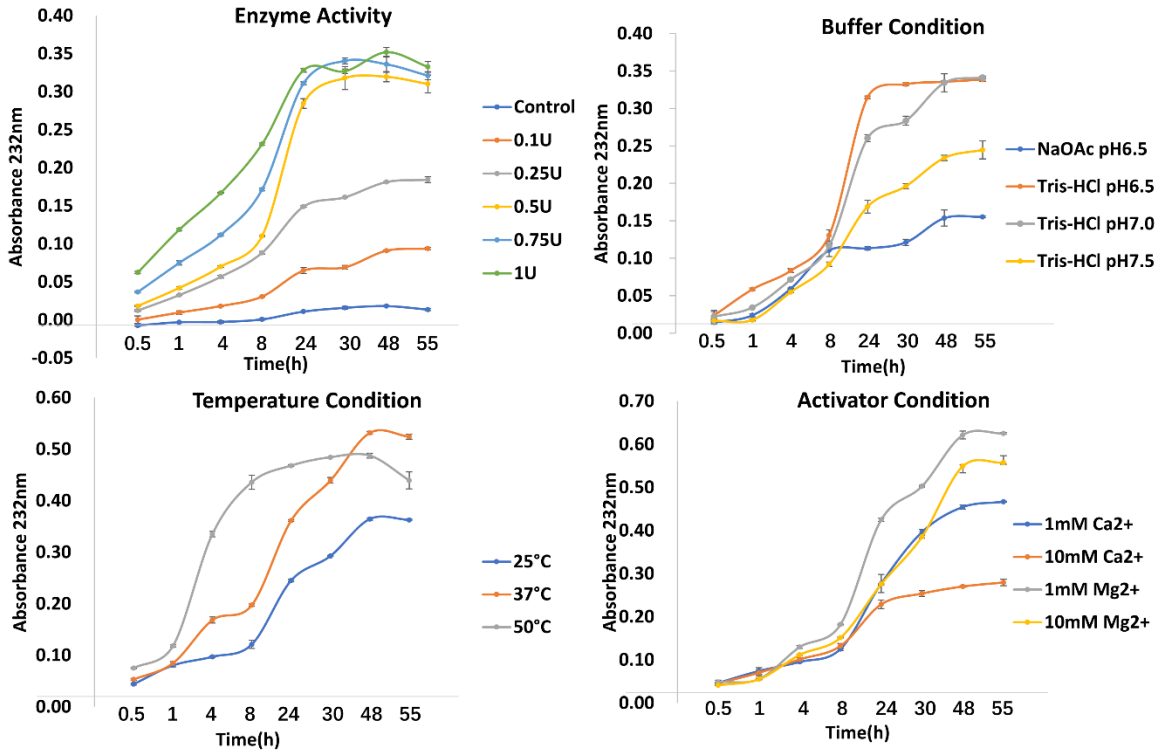


Figure 4.4 Heparinase digestion of Pn8P reaction condition optimization. A. Enzyme amount optimization; B. buffer condition optimization; C. Temperature condition optimization; D. activators condition optimization.

4.4.3 Separation and analysis of digested product

Although the HPLC-ELSD method could monitor the Pn8P reaction quickly, the short column used in this method could not separate different oligosaccharide because of its low resolution capabilities in the low molecular weight range. The ELSD detector does not allow for recovery of the sample injected, so the eluted oligosaccharide could not be further analyzed by Mass Spectrometry. Therefore, a longer SEC column better suited for the smaller oligosaccharides, coupled with HPLC-RID was applied for further analysis. From the RID results, the Pn8P was depolymerized to trisaccharide, tetrasaccharide, octasaccharide and high molecular weight oligosaccharides. (Figure 4.8)

The low molecular weight oligosaccharides were confirmed by Fourier Transform based Mass Spectrometry (FTMS), and the error was calibrated to be within 5ppm. The octasaccharide and tetrasaccharide were detected under negative mode, while the trisaccharide was detected under positive mode. The mass to charge ratio of dp8, tetrasaccharide and trisaccharide are 1323.3710 m/z, 661.1809 m/z and 504.1877 m/z respectively. (Table 4.1, Figure 4.5, 4.9, 4.10)

Table 4.1 Low molecular Oligosaccharides detected by FTMS

	m/z	charge	Theoretical Mass	Error(ppm)
Octasaccharide	1323.3710	1	1324.3811	1.66
Tetrasaccharide	661.1809	1	662.1906	2.72
Trisaccharide	504.1877	1	503.1612	3.71

4.4.4 Characterization of oligosaccharide by MSn

The repeat unit of Pn8P, the tetrasaccharide was further studied in multiple stage tandem mass spectrometry (MSn) mode by MS. (Figure 4.5) The peak at 661.1809 m/z was selected as the precursor ion, then sequential fragmentation producing ions 661→499→337→158 gave a complete sequence of the sugars in the tetrasaccharide repeat unit. In the MS spectrum, the fragment ions were mainly the trisaccharide ion (C3) and the disaccharide ion (C2) from the non-reducing end with some cross ring internal cleavage on glucuronic acid (1,3X4) and the galactose (2,4A5, 2,5A5). (Figure 4.5, MS2 spectrum) In the MS3 spectrum, the C3 ion was further dissociated into another disaccharide ion (B2) and Glucuronic acid ion (C1). (Figure 4.5, MS3 spectrum) In the MS4 spectrum, the acquisition of the B1 ion confirming the unsaturated C-C bond on glucuronic acid, and more fragment ions with 1,3X4 ion further confirm the previous results.(Figure 4.5, MS4 spectrum)

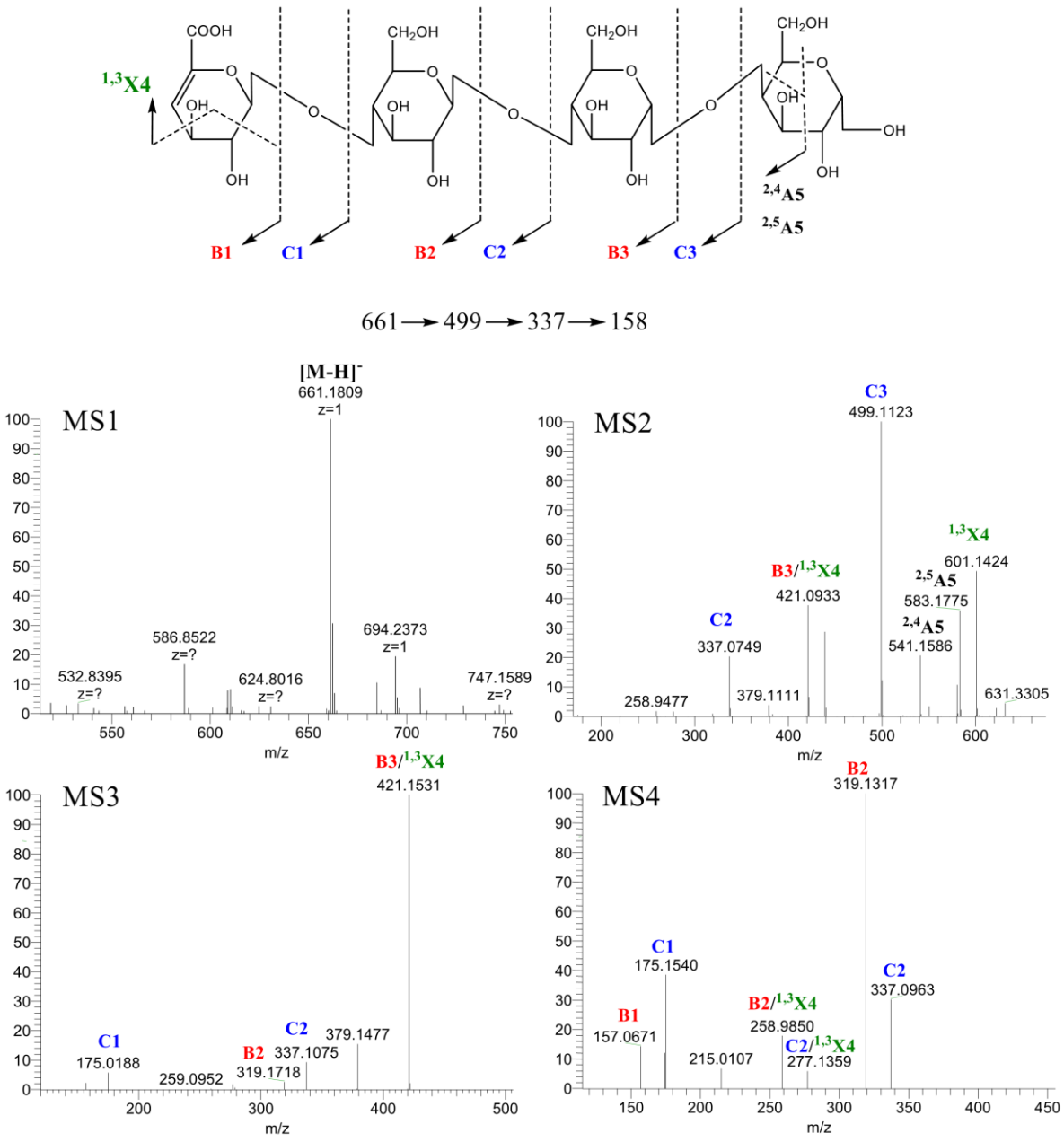


Figure 4.5 FTMS and MS4 results of tetrasaccharide from HPLC-SEC-RID.

4.4.5 Characterization of oligosaccharide by NMR

In order to further characterize the digested polysaccharide and confirm the enzymatic cleavage at the α-linkage between galactose and glucuronic acid moiety, a set of 2D homo- and heteronuclear NMR experiments were performed. (See in SI, NMR spectroscopy section) Here, we compared spectra of the polysaccharide before and after digestion (Figure 4.6). The

assignments were based on the reported chemical shift of native Pn8P10. As shown in figure 4.6, both C1 of Glc C and B1 of Glc B have similar chemical shifts after digestion, and only minor proton shifts were observed in the heteronuclear single quantum coherence (HSQC) spectra (less than 0.05 ppm). However, a chemical shift change was seen in the glucuronic acid's D1 signal from H=4.56 ppm, C=102.3 ppm to D'1 H=5.13 ppm, C=100.5 ppm. Concomitantly, an alkene peak at H=5.85 ppm, C=110.1 ppm) was observed and determined to be the proton from the D'4. Strong correlation was observed in the TOCSY experiment between D'1 and D'4 (supporting information) as further evidence that the D'4 ring is hydrolyzed during digestion. Based on the LCMS data, it is anticipated that the α glycosylic linkage of A1 of galactose A is hydrolyzed and the digestion will result in an α/β mixture. The signals for the α and β configuration (~1:2 ratio) belonging to the A'1 were observed in the HSQC spectra, which was further verified by the well resolved α and β multiplet patterns in the t2-coupled HSQC experiment 24(supporting information). This evidence supports the previous finding that b-elimination happens at the glucuronic acid moiety and the enzyme mixture promotes α-linkage.

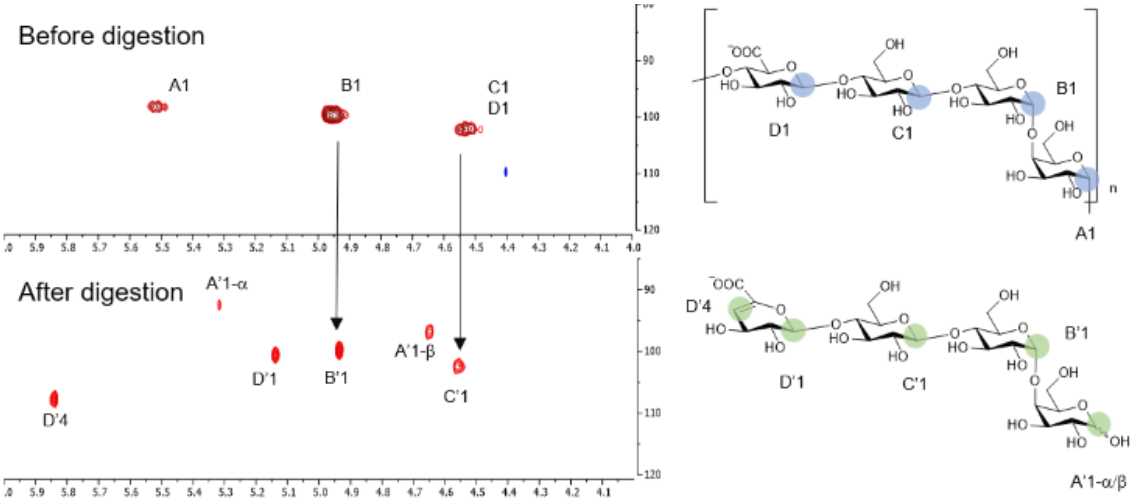


Figure 4.6 HSQC spectra of before/after digestion of Pn8P.

4.5 Conclusion

In this study, the heparinase from Flavobacterium heparinum has been applied to successfully depolymerize pneumococcal type 8 polysaccharide, and the oligosaccharide repeat unit of the polysaccharide has been characterized by HPLC, MS4 and NMR approaches. The results indicate that heparinase is able to cleave α glycosidic bond between galactose and glucuronic acid. These results open a new direction that GAG degrading enzyme can be used to depolymerize and study pneumococcal polysaccharide. Although the commercially available heparinase I&III demonstrated reproducibility between different labs, it is still difficult to rule out the possibility that the contamination of the other proteins may happen during the enzyme production.

4.6 Reference

1. Kenneth, J. R.; Ryan, C.; Ray, G., Sherris medical microbiology. *Vasa* **2004**.
2. Geno, K. A.; Gilbert, G. L.; Song, J. Y.; Skovsted, I. C.; Klugman, K. P.; Jones, C.; Konradsen, H. B.; Nahm, M. H., Pneumococcal capsules and their types: past, present, and future. *Clin. Microbiol. Rev.* **2015,** *28* (3), 871-899.
3. Avci, F. Y.; Kasper, D. L., How bacterial carbohydrates influence the adaptive immune system. *Annu. Rev. Immunol.* **2009,** *28*, 107-130.
4. Hill, A. B.; Beitelshees, M.; Nayerhoda, R.; Pfeifer, B. A.; Jones, C. H. J. A. i. d., Engineering a next-generation glycoconjugate-like Streptococcus pneumoniae vaccine. **2018,** *4* (11), 1553-1563.
5. KAMERLING, J. P., Pneumococcal polysaccharides: a chemical view. *Streptococcus pneumoniae: molecular biology and mechanisms of disease* **1999**, 81-114.
6. Feng, S.; Xiong, C.; Wang, S.; Guo, Z.; Gu, G. J. A. i. d., Semi-synthetic Glycoconjugate Vaccines to Elicit T Cell-Mediated Immune Responses and Protection against Streptococcus pneumoniae Serotype 3. **2019**.
7. Magee, A. D.; Yother, J., Requirement for Capsule in Colonization byStreptococcus pneumoniae. *Infect. Immun.* **2001,** *69* (6), 3755-3761.
8. Morrison, K. E.; Lake, D.; Crook, J.; Carlone, G. M.; Ades, E.; Facklam, R.; Sampson, J. S., Confirmation of psaA in All 90 Serotypes ofStreptococcus pneumoniae by PCR and Potential of This Assay for Identification and Diagnosis. *J. Clin. Microbiol* **2000,** *38* (1), 434-437.
9. Bentley, S. D.; Aanensen, D. M.; Mavroidi, A.; Saunders, D.; Rabbinowitsch, E.; Collins, M.; Donohoe, K.; Harris, D.; Murphy, L.; Quail, M. A., Genetic analysis of the capsular biosynthetic locus from all 90 pneumococcal serotypes. *PLoS. Genet.* **2006,** *2* (3), e31.
10. Schumann, B.; Hahm, H. S.; Parameswarappa, S. G.; Reppe, K.; Wahlbrink, A.; Govindan, S.; Kaplonek, P.; Pirofski, L.-a.; Witzenrath, M.; Anish, C., A semisynthetic

Streptococcus pneumoniae serotype 8 glycoconjugate vaccine. *Sci. Transl. Med.* **2017,** *9* (380), eaaf5347.
11. Gaborieau, M.; Castignolles, P., Size-exclusion chromatography (SEC) of branched polymers and polysaccharides. *Anal. Bionanl. Chem.* **2011,** *399* (4), 1413-1423.
12. Mori, S.; Barth, H. G., *Size exclusion chromatography*. Springer Science & Business Media: 2013.
13. Hadidi, M.; Buckley, J. J.; Zydney, A. L., Effects of solution conditions on characteristics and size exclusion chromatography of pneumococcal polysaccharides and conjugate vaccines. *Carbohydr. Polym.* **2016,** *152*, 12-18.
14. Gairola, S.; Gautam, M.; Patil, D.; Kumar, K. M.; Shinde, P.; Jana, S.; Dhere, R.; Jadhav, S., Development and application of HPLC-RI and HPLC-MS/MS based methods for quantification of residual deoxycholate levels in pneumococcal polysaccharides. *Biol.* **2016,** *44* (6), 517-525.
15. Xu, J.; Guan, J.; Chen, X.; Zhao, J.; Li, S., Comparison of polysaccharides from different Dendrobium using saccharide mapping. *J. Pharm. Biomed. Anal.* **2011,** *55* (5), 977-983.
16. Li, L.; Zhang, F.; Zaia, J.; Linhardt, R. J., Top-down approach for the direct characterization of low molecular weight heparins using LC-FT-MS. *Anal. Chem.* **2012,** *84* (20), 8822-8829.
17. Wang, Z.; Zhang, T.; Xie, S.; Liu, X.; Li, H.; Linhardt, R. J.; Chi, L., Sequencing the oligosaccharide pool in the low molecular weight heparin dalteparin with offline HPLC and ESI–MS/MS. *Carbohydr. Polym.* **2018,** *183*, 81-90.
18. Zhang, T.; Liu, X.; Li, H.; Wang, Z.; Chi, L.; Li, J.-P.; Tan, T., Characterization of epimerization and composition of heparin and dalteparin using a UHPLC-ESI-MS/MS method. *Carbohydr. Polym.* **2019,** *203*, 87-94.
19. Li, L.; Li, Y.; Ijaz, M.; Shahbaz, M.; Lian, Q.; Wang, F., Review on complement analysis method and the roles of glycosaminoglycans in the complement system. *Carbohydr. Polym.* **2015,** *134*, 590-597.
20. Wang, W.; Wang, J.; Li, F., Hyaluronidase and chondroitinase. In *Protein Reviews*, Springer: 2016; pp 75-87.
21. Garron, M.-L.; Cygler, M., Structural and mechanistic classification of uronic acid-containing polysaccharide lyases. *Glycobiology* **2010,** *20* (12), 1547-1573.
22. Ernst, S.; Langer, R.; Cooney, C. L.; Sasisekharan, R., Enzymatic degradation of glycosaminogIycans. *Crit. Rev. Biochem. Mol. Biol.* **1995,** *30* (5), 387-444.
23. Middleton, D. R.; Zhang, X.; Wantuch, P. L.; Ozdilek, A.; Liu, X.; LoPilato, R.; Gangasani, N.; Bridger, R.; Wells, L.; Linhardt, R. J., Identification and characterization of the Streptococcus pneumoniae type 3 capsule-specific glycoside hydrolase of Paenibacillus sp ecies 32352. *Glycobiology* **2017,** *28* (2), 90-99.

4.7 Appendix

Table 4.2 Optimization of heparinase amount required for reaction by nanodrop

Enzyme Amount	30min	1h	4h	8h	24h	30h	48h	55h
control	-0.0080	-0.0020	-0.0040	0.0000	0.0110	0.0150	0.0180	0.0120
	-0.0080	-0.0040	-0.0030	0.0000	0.0110	0.0150	0.0190	0.0150
	-0.0060	-0.0030	-0.0010	0.0020	0.0110	0.0180	0.0180	0.0140
0.1U	0.0060	0.0110	0.0185	0.0300	0.0675	0.0700	0.0900	0.0930
	-0.0020	0.0070	0.0175	0.0315	0.0665	0.0665	0.0920	0.0955
	-0.0030	0.0110	0.0185	0.0305	0.0605	0.0710	0.0910	0.0930
0.25U	0.0130	0.0330	0.0550	0.0865	0.1500	0.1620	0.1820	0.1860
	0.0110	0.0320	0.0565	0.0900	0.1480	0.1610	0.1810	0.1800
	0.0130	0.0330	0.0585	0.0885	0.1490	0.1610	0.1810	0.1870
0.5U	0.0190	0.0400	0.0700	0.1095	0.2790	0.3040	0.3120	0.3231
	0.0180	0.0430	0.0710	0.1110	0.2830	0.3340	0.3230	0.3000
	0.0180	0.0430	0.0690	0.1105	0.2915	0.3160	0.3240	0.3075
0.75U	0.0360	0.0780	0.1115	0.1705	0.3090	0.3385	0.3300	0.3265
	0.0380	0.0750	0.1105	0.1705	0.3110	0.3380	0.3480	0.3205
	0.0360	0.0720	0.1130	0.1735	0.3130	0.3450	0.3300	0.3160
1U	0.0640	0.1190	0.1670	0.2295	0.3275	0.3240	0.3555	0.3320
	0.0620	0.1170	0.1665	0.2320	0.3260	0.3265	0.3445	0.3260
	0.0610	0.1200	0.1670	0.2320	0.3310	0.3300	0.3555	0.3400

Table 4.3 Optimization of heparinase activity in different temperature by nanodrop

Temperature	30min	1h	4h	8h	24h	30h	48h	55h
25°C	0.0430	0.0820	0.0990	0.1120	0.2470	0.2940	0.3670	0.3640
	0.0460	0.0820	0.0950	0.1260	0.2450	0.2920	0.3630	0.3620
	0.0430	0.0770	0.0970	0.1250	0.2430	0.2920	0.3630	0.3620
37°C	0.0530	0.0880	0.1660	0.1960	0.3630	0.4440	0.5320	0.5220
	0.0550	0.0840	0.1640	0.2000	0.3600	0.4410	0.5290	0.5300
	0.0510	0.0820	0.1750	0.1960	0.3610	0.4330	0.5340	0.5210
50°C	0.0740	0.1200	0.3290	0.4300	0.4700	0.4840	0.4840	0.4580
	0.0760	0.1190	0.3370	0.4260	0.4670	0.4840	0.4850	0.4330
	0.0750	0.1160	0.3390	0.4510	0.4670	0.4850	0.4930	0.4270

Table 4.4 Optimization of heparinase activity in different buffer condition by nanodrop

Buffer	30min	1h	4h	8h	24h	30h	48h	55h
Sodium Acetate pH 6.5	0.0120	0.0220	0.0590	0.1010	0.1110	0.1180	0.1660	0.1560
	0.0140	0.0260	0.0600	0.1160	0.1130	0.1200	0.1460	0.1560
	0.0170	0.0240	0.0600	0.1150	0.1160	0.1260	0.1490	0.1540
Tris-HCl pH 6.5	0.0290	0.0570	0.0810	0.1220	0.3140	0.3330	0.3340	0.3400
	0.0190	0.0590	0.0840	0.1360	0.3140	0.3300	0.3370	0.3360
	0.0230	0.0600	0.0860	0.1340	0.3170	0.3340	0.3360	0.3400
Tris-HCl pH 7.0	0.0250	0.0350	0.0700	0.1120	0.2630	0.2780	0.3250	0.3420
	0.0290	0.0340	0.0720	0.1180	0.2550	0.2830	0.3300	0.3390
	0.0120	0.0340	0.0730	0.1200	0.2630	0.2900	0.3480	0.3430
Tris-HCl pH 7.5	0.0170	0.0160	0.0560	0.0920	0.1610	0.2000	0.2380	0.2310
	0.0180	0.0190	0.0540	0.0890	0.1680	0.1940	0.2320	0.2520
	0.0150	0.0180	0.0550	0.0950	0.1780	0.1950	0.2320	0.2510

Table 4.5 Optimization of heparinase activity in different activator condition by nanodrop

Activator	30min	1h	4h	8h	24h	30h	48h	55h
1mM $CaCl_2$	0.0450	0.0660	0.0940	0.1220	0.2770	0.4000	0.4490	0.4660
	0.0460	0.0750	0.0950	0.1240	0.2980	0.3890	0.4550	0.4660
	0.0480	0.0810	0.0960	0.1290	0.2560	0.4000	0.4590	0.4680
10Mm $CaCl_2$	0.0460	0.0640	0.0970	0.1280	0.2400	0.2540	0.2720	0.2880
	0.0440	0.0690	0.1040	0.1390	0.2200	0.2470	0.2690	0.2760
	0.0470	0.0770	0.1020	0.1290	0.2260	0.2600	0.2690	0.2740
1mM MgCl2	0.0440	0.0560	0.1330	0.1830	0.4200	0.5020	0.6120	0.6270
	0.0510	0.0570	0.1290	0.1810	0.4270	0.5010	0.6300	0.6250
	0.0510	0.0530	0.1270	0.1820	0.4280	0.5050	0.6220	0.6230
10mM MgCl2	0.0410	0.0550	0.1130	0.1510	0.2760	0.3890	0.5330	0.5530
	0.0420	0.0540	0.1090	0.1510	0.2730	0.3860	0.5500	0.5610
	0.0400	0.0560	0.1130	0.1520	0.2780	0.3820	0.5640	0.5590

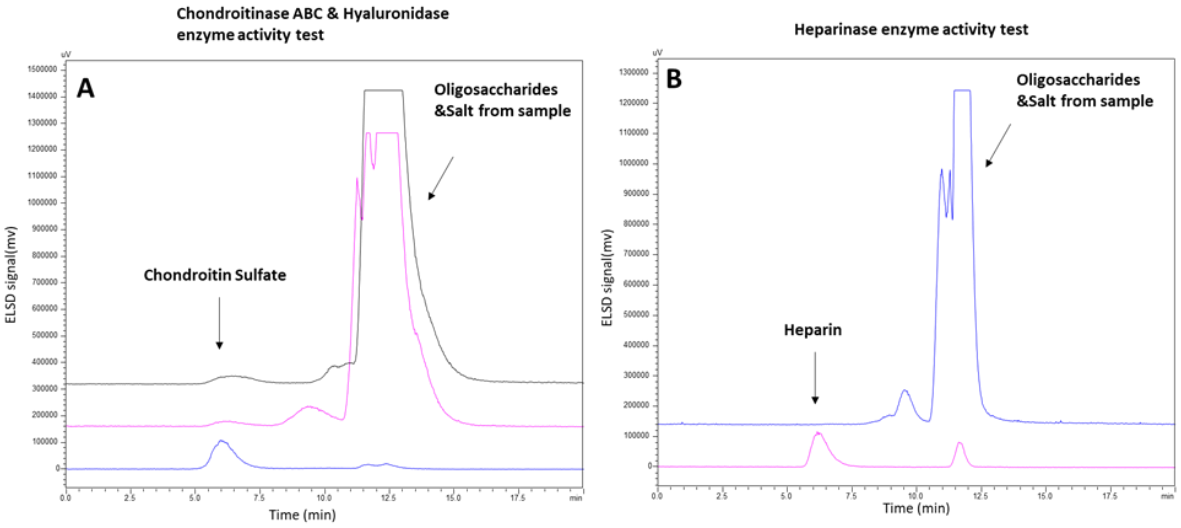


Figure 4.7 SEC-ELSD results of enzyme activity tests. A. Chondroitinase ABC reaction with chondroitin sulfate (dark), hyaluronidase reaction with chondroitin sulfate (red), Chondroitin Sulfate control without enzyme (blue), the results indicate that all the enzyme had ideal activity; B. Heparinase I&III reaction with heparin (blue), heparin control without enzyme (red).

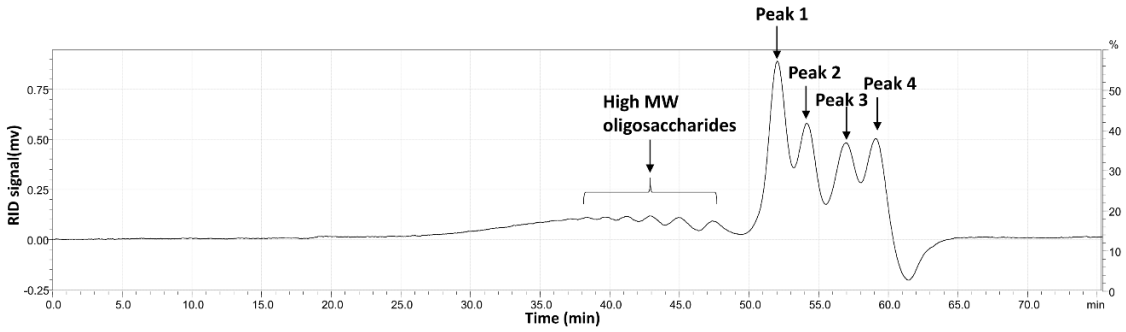


Figure 4.8 HPLC results of digested Pn8P by SEC column separation and RID detection, the Pn8P was depolymerized to trisaccharide (peak3), tetrasaccharide (peak2), octasaccharide (peak1) and high molecular weight oligosaccharides.

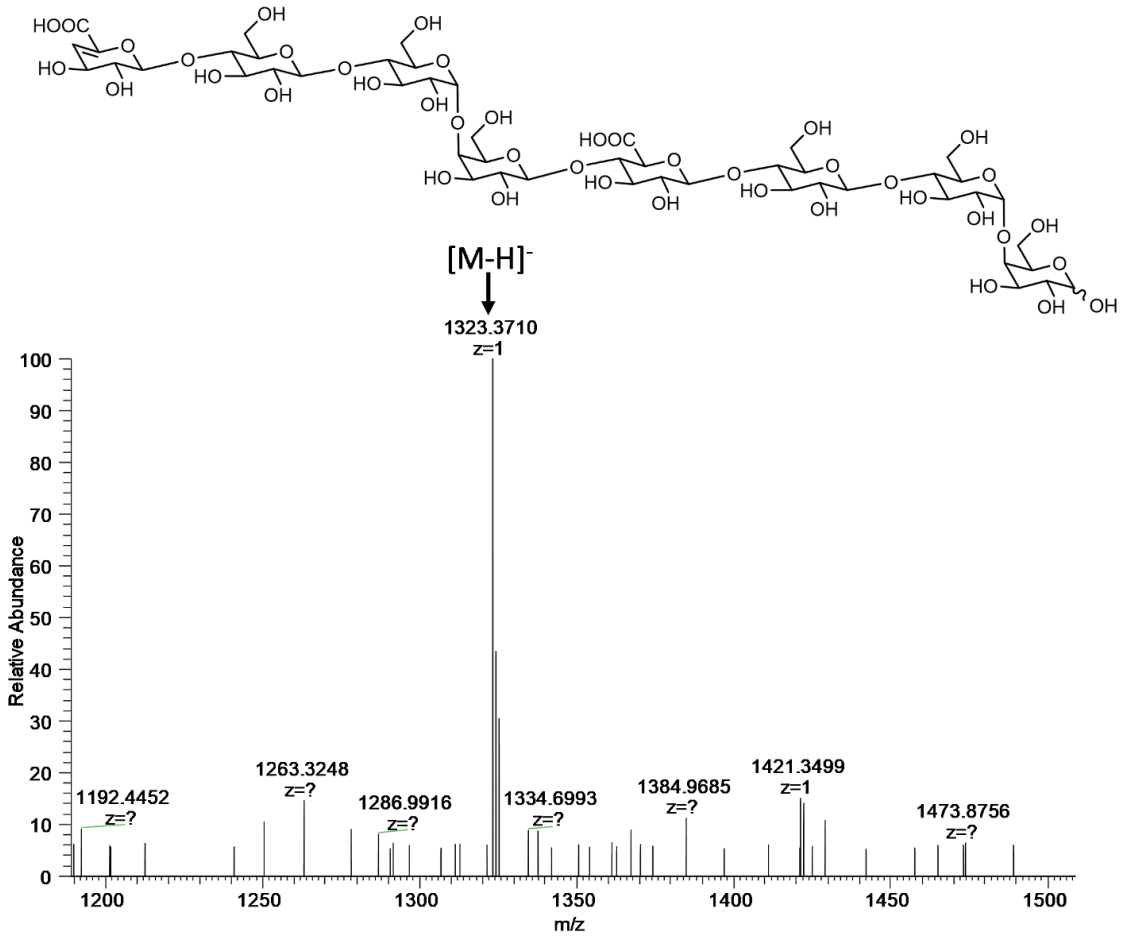


Figure 4.9 FTMS spectra of SEC-RID peak 1 from HPLC, octasaccharide.

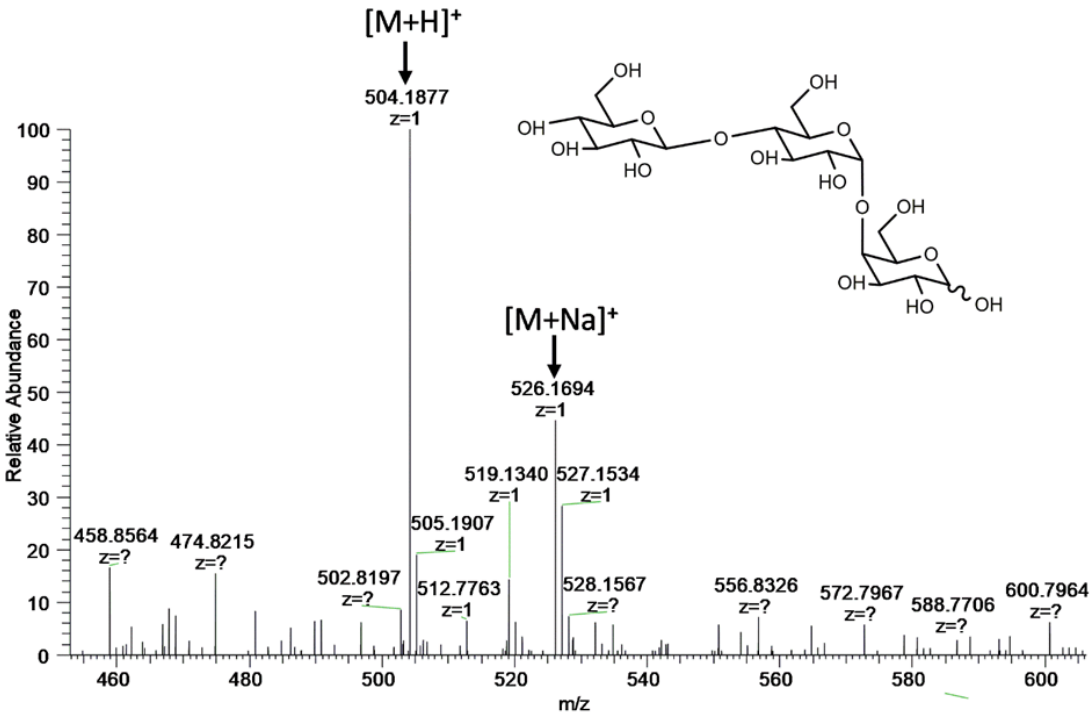


Figure 4.10 FTMS spectra of SEC-RID peak 3 from HPLC, trisaccharide.

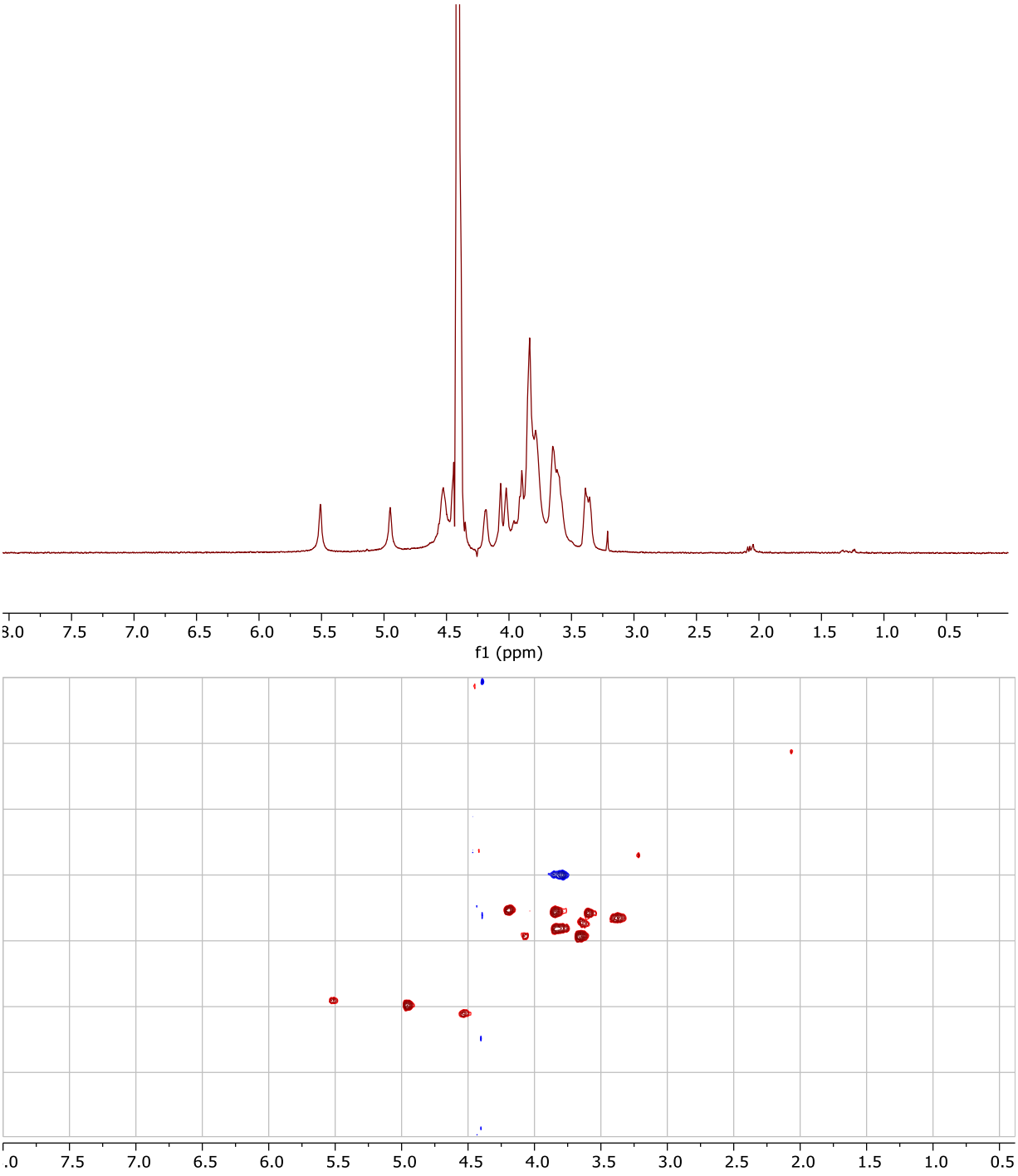


Figure 4.11 Proton and HSQC data for Pn8 polysaccharide.

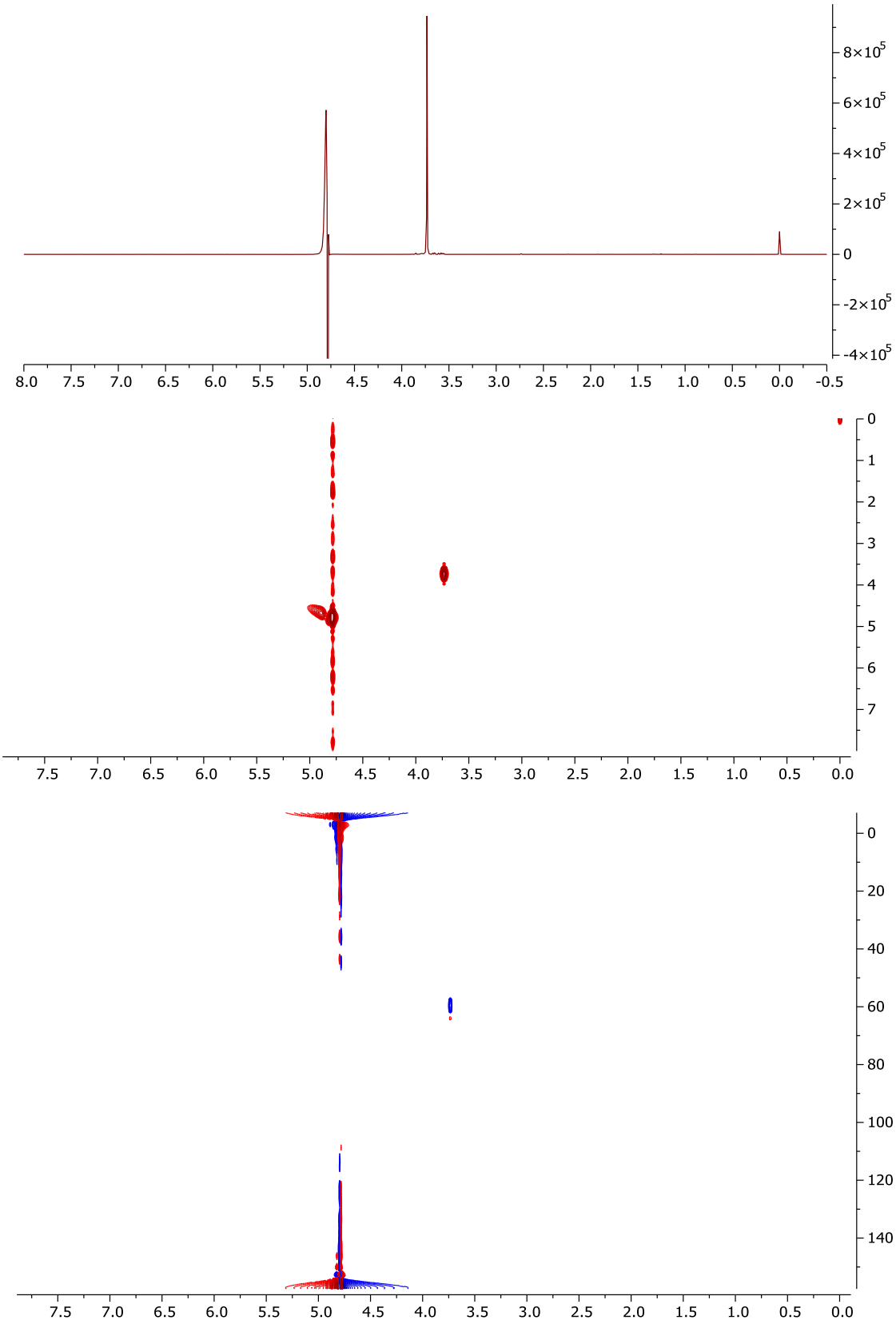


Figure 4.12 Proton, COSY and SDQC data for digestion blank (buffer + enzyme).

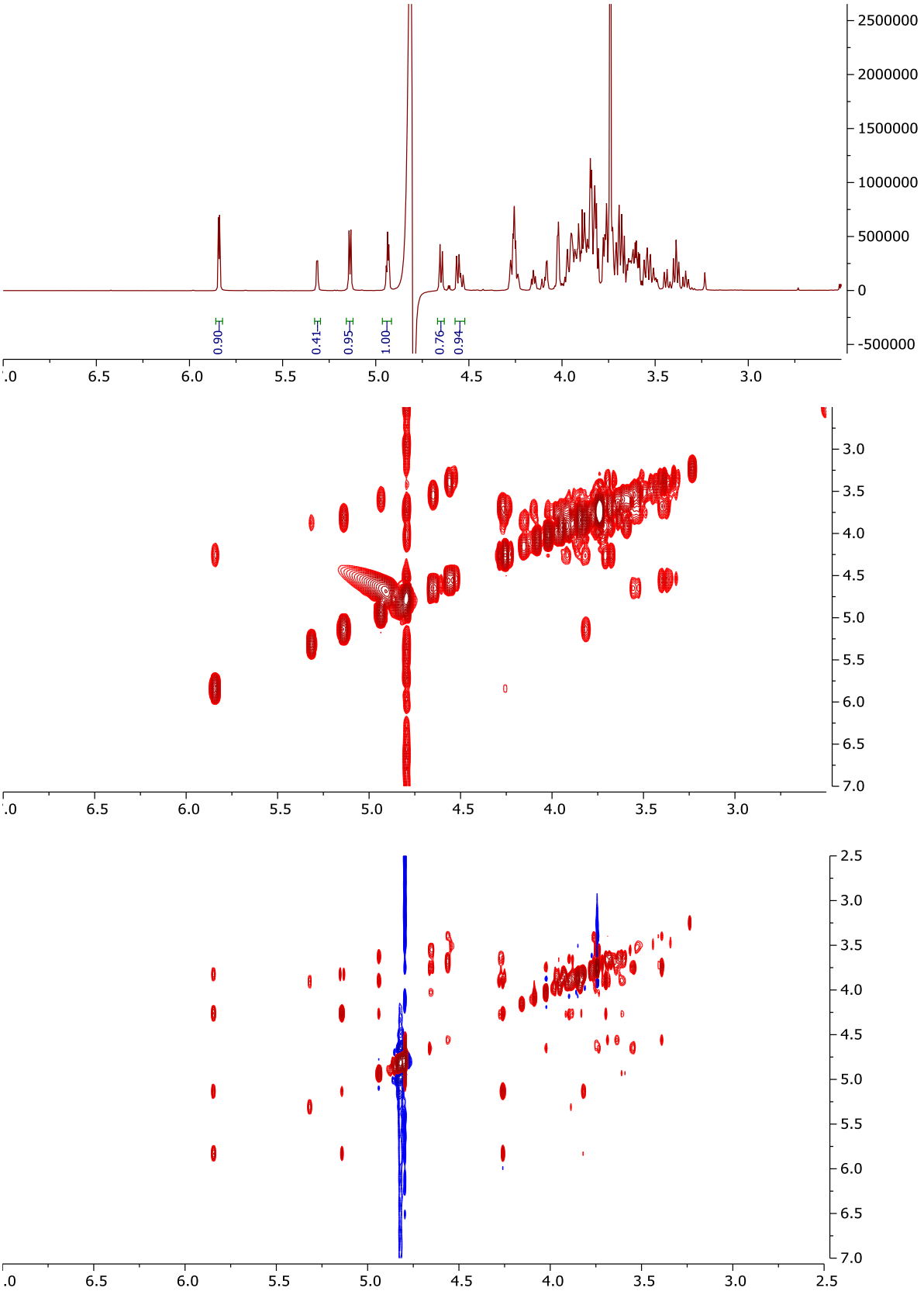


Figure 4.13 Proton, TOCSY and decoupled HSQC data for digested polysaccharide.

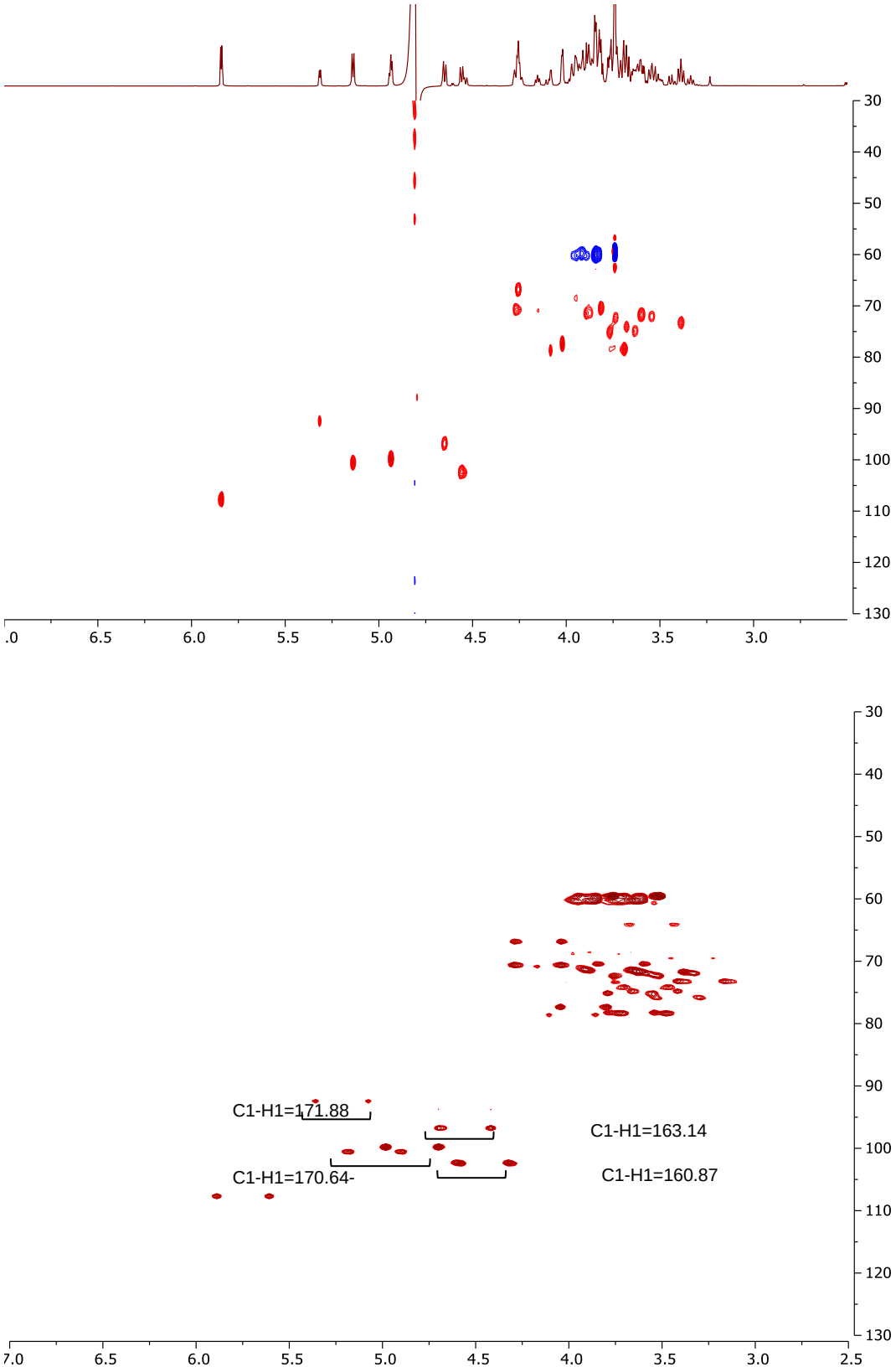


Figure 4.14 COSY, HSQC data for digested polysaccharide.