

A Procedure for the Analysis of Site-specific and Structure-specific

Fucosylation in Alpha-1-Antitrypsin

Haidi Yin^{1,2}, Jianhui Zhu¹, Jing Wu¹, Zhijing Tan¹, Mingrui An¹, Shiyue Zhou³, Yehia Mechref³ and

David M. Lubman^{1*}

¹Department of Surgery, University of Michigan Medical Center, Ann Arbor, MI 48109

²Department of Applied Biology and Chemical Technology, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

³Department of Chemistry and Biochemistry, Texas Tech University, Lubbock, TX 79409

Author Email Address

HaidiYin haidi.yin@polyu.edu.hk

Jianhui Zhu jianhuiz@umich.edu

Jing Wu jingwu@umich.edu

Zhijing Tan zhijing@umich.edu

Mingrui An mingruia@umich.edu

Shiyue Zhou shiyue.zhou@ttu.edu

Yehia Mechref yehia.mechref@ttu.edu

David M. Lubman dmlubman@umich.edu

Running Title: site-specific and structure-specific fucosylation analysis of alpha-1-Antitrypsin

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1002/elps.201600176](https://doi.org/10.1002/elps.201600176).

This article is protected by copyright. All rights reserved.

* To whom correspondence should be addressed. David M. Lubman, Department of Surgery, University of Michigan Medical Center, 1150 West Medical Center Drive, Building MSRB1 Rm A510B, Ann Arbor, MI 48109-0656. E-mail: dmlubman@umich.edu; Phone: 734-647-8834; Fax: 734-615-2088

Abbreviations: A1AT, alpha-1-antitrypsin; LC-MS/MS, liquid chromatography with tandem mass spectrometry; TCEP, tris (2-carboxyethyl) phosphine, IAA, iodoacetamide; TEAB, Triethylammonium bicarbonate; TFA, trifluoroacetic acid; HILIC, hydrophilic interaction liquid chromatography

Keywords: alpha-1-antitrypsin, antennary fucosylation, core fucosylation, glycan structure, liquid chromatography-mass spectrometry

Total Number of Words: 4419

Abstract

A mass spectrometry-based methodology has been developed for analysis of core-fucosylated versus antennary-fucosylated glycosites in glycoproteins. This procedure is applied to the glycoprotein alpha-1-antitrypsin (A1AT) which contains both core- and antennary-fucosylated glycosites. The workflow involves digestion of intact glycoproteins into glycopeptides, followed by double digestion with sialidase and galactosidase. The resulting glycopeptides with truncated glycans were separated using an off-line HILIC (Hydrophilic Interaction Liquid Chromatography) separation where multiple fractions were collected at various time intervals. The glycopeptides in each fraction were treated with PNGase F and then divided into halves. One half of the sample was applied for peptide identification while the other half was processed for glycan analysis by derivatizing with a Meladrazine reagent followed by mass spectrometry analysis. This procedure provided site-specific identification of glycosylation sites and the ability to distinguish core fucosylation and antennary

fucosylation via a double digestion and a mass profile scan. Both core and antennary fucosylation are shown to be present on various glycosites in A1AT.

1 Introduction:

Recent studies have shown that protein glycosylation can be related to disease states[1, 2]. In particular, fucosylation has been studied as a potential marker for several different cancers and diseases, such as hepatocellular carcinoma[3], pancreatic cancer[4] and lung cancer[5]. In some studies total fucosylation has often been shown to be related to disease state by using lectin blots[6-8]. However, more recently it has been found that specific fucosylation structures and the specific fucosylation sites of a glycoprotein may be more directly related to changes in disease[9, 10].

The fucosylation structures include core fucosylation and antennary fucosylation where both may be present on any glycosylation site. One fucosylated glycoprotein usually has several fucosylation sites. Lectin blots can neither provide the fucosylation level of each site of the proteins nor distinguish the two fucosylation structures specifically. Some studies have used lectins to enrich the core-fucosylated or antennary-fucosylated glycoproteins before removing the glycan and analyze proteins and/or glycans separately[11, 12], where for example *Aleuria Aurantia* lectin (AAL) has high affinity to both core- and antennary-fucosylated glycoproteins whereas *Lens culinaris* lectin (LCA) has high affinity

to core-fucosylated glycoproteins. However, such analyses missed the site-specific information and cannot provide direct evidence for core fucosylation or antennary fucosylation of these glycoproteins.

Mass spectrometry (MS) can be used to identify most structures of glycans[13-15], but it is often difficult to do so for fucosylated glycans because the fucose glycosidic bond is labile in dissociation where even MS/MS may not distinguish whether it is core- or antennary-fucosylated. A widely accepted glycan structure analysis approach to distinguish the fucosylation structures uses various fucosidases to digest the glycan and analyze the glycan several times with HPLC (high pressure liquid chromatography), but this approach is very time consuming, which usually involves various fucosidases and several runs of HPLC analysis[5]. Recent studies have analyzed the glycopeptide with truncation of the glycan at the core glycan structure between GlcNAC-GlcNAC (N-acetylglucosamine)[16, 17]. A core-fucosylated glycopeptide with core-fucose modification on the inner GlcNAC could be easily distinguished with mass spectrometry, where site-specific information was also maintained. However, this procedure cannot be applied to antennary fucosylation structures.

In current work we have developed a methodology to distinguish core fucosylation and antennary fucosylation of glycans at each site of a glycoprotein as applied to the protein A1AT. The procedure involves a double digestion of glycopeptides with sialidase and galactosidase, followed by an off-line HILIC (Hydrophilic Interaction Liquid Chromatography) column separation of glycopeptides with a truncated glycan. Using this method only a single run of LC-MS analysis is needed for the sialidase/galactosidase digestion procedure to differentiate core/antennary fucosylation, whereas at least two runs are needed for the fucosidase digestion procedure, one run without fucosidase digestion and another run with corresponding fucosidase digestion. The glycopeptides in each fraction were further digested with PNGase F where the peptides are analyzed by LC-MS/MS (liquid chromatography with tandem mass spectrometry) directly and the glycans are derivatized by Meladrazine reagent before LC-MS analysis. This procedure has been used to identify site-specific

glycosylation and to distinguish core fucosylation and antennary fucosylation via use of the mass profile scan in the protein A1AT.

2 Materials and Methods

2.1 Trypsin Digestion, Followed by Sialidase/Galactosidase Double Digestion: For trypsin digestion we added 10 μL of 100 mM triethylammonium bicarbonate (TEAB) to 30 μg alpha-1-antitrypsin (A1AT) and pipetted to dissolve the sample well. Then dissolved A1AT was reduced with 10 mM tris (2-carboxyethyl) phosphine (TCEP) at 37°C for 30 min and alkylated with 20 mM iodoacetamide (IAA) at room temperature in the dark for 15 min. The sample solution is diluted for 3 times with 100 mM TEAB and incubated with 2 μL (0.5 $\mu\text{g}/\mu\text{L}$) trypsin (Promega, Madison, WI) at 37°C for 16 h. The trypsin is eventually deactivated at 95°C for 5 min and dried in a speedvac.

For the sialidase and galactosidase digestion, the glycopeptide mixture was dissolved in 30 μL of 25 mM sodium acetate solution (pH5.5) and incubated with 15 mU (3 μL) of non-specific sialidase and 75 mU (3 μL) of β (1-4,6) galactosidase (Prozyme, Hayward, CA) at 37°C for 18 h to remove all sialic acid residues and remove galactose on condition that no fucose is bound to the subterminal N-acetylglucosamine in an N-glycan. The glycosidases were deactivated at 95°C for 5 min. In order to confirm core and antennary fucosylation of the sialidase/galactosidase double digested glycans, various fucosidases were used to cut the corresponding fucose as described by Varadi et al.[18].

2.2 C18 Desalting: Trifluoroacetic acid (TFA) was added until the pH value reached 2. The columns were pre-wetted with 200 μL 0.1% TFA in 50% acetonitrile for 5 times and equilibrated with 0.1% TFA in water for 3 times by centrifugation at 1,500 g/min for 1 min each time. The peptides were bound to the C18 medium for 5 times followed by 3 times washing with 0.1% TFA to remove non-

specific binding by centrifugation as described above; 20 μ L of 50% acetonitrile with 0.1% TFA was used for elution by centrifugation as described above. Elution was repeated once and the combined eluents were then dried in a speedvac.

2.3 Off-line HILIC Fractionation: The column used was a Waters Xbridge HILIC column, 2.1 mm \times 150 mm, 5 μ m. The separations were performed on a Beckman HPLC system with detection at 214 nm. The sample was diluted to 100 μ L with initial phase, and filtered with a 0.45 μ m membrane. The gradient was set from 80% to 50% Solvent B (Solvent A: 10 mM ammonium formate; Solvent B: 980 ml Acetonitrile, 20 ml H₂O, 0.1% formic acid) with a flow rate 0.2 ml/min and collection every 2 min.

2.4 PNGase F digestion, Purification of Deglycosylated Peptides and Extraction of Glycan: N-Glycosidase F (PNGase F, 2.5 U/10 μ g protein) (New England Biolabs, Ipswich, MA) was added and incubated at 37°C for 18 h. PNGase F was deactivated by boiling for 5 min. Half of the protein glycan mixture was dried in a SpeedVac, desalted with a C18 column as described above and dried for further peptide identification by LC-MS. The other half was used for glycan structure analysis.

10 μ L porous graphitized carbon tips (PGC tips) (Sigma Aldrich, St. Louis, MO) were used to separate glycans from proteins and other impurities. The tip was activated by 50% ACN (with 0.1% TFA) and equilibrated with water (with 0.1% TFA). The samples were then loaded to the tip followed by 3 washes with water (with 0.1% TFA) to remove nonspecific binding. Then 10 μ L of 10% ACN (with 0.1% TFA) and 10 μ L of 25% ACN (with 0.1% TFA) were used for glycan elution where the two eluents were combined and dried down for further glycan analysis.

2.5 LC-MS Identification of Peptides: Nano LC-MS/MS conditions were as described in previous work[17]. A Magic C18 capillary column (100 μ m \times 15 cm; 3 μ m particles, 200 Å) (Michrom Biosciences, Auburn, CA) was used for LC separation, and gradient elution was performed using a Paradigm MG4 micro-pump system (Michrom Biosciences, Auburn, CA) with a flow rate of 350

nL/min. The mobile phase A was 2% acetonitrile with 0.1% formic acid in water and mobile phase B was 2% water with 0.1% formic acid in acetonitrile. The analytical gradient lasted for 80 min where the composition of solvent B rose from 5% to 32% in 50 min, followed by a washing and equilibration step where solvent B increased to 95% in 1 min and was held for 4 min, and then returned to 5% B in 0.1 min and was held for 25 min.

An ESI-LTQ mass spectrometer (Thermo Fisher Scientific, San Jose, CA) operated in positive ion mode was used for analysis. The ESI spray voltage and capillary voltage were set at 2.2 kV and 45 V respectively. Collision ionization dissociation (CID) fragmentation was performed at 35% of the normalized collision energy. The mass spectra were acquired in a data-dependent manner. Following a full scan in the mass range of m/z 400 to 1800, CID MS/MS was performed on the five most intense ions.

2.6 Database Search for Peptide Identification: The mass spectrum was searched with Proteome Discoverer 1.2 (Thermo Fisher Scientific, San Jose, CA) software with SEQUEST using the following settings: (1) fixed modification: cysteine carbamidomethylation (+57.0 Da); (2) Dynamic modification: methionine oxidation (+16.0 Da), addition of +1 Da to Asparagine residue due to PNGase digestion; (3) One missed cleavage was allowed; (4) Peptide ion tolerance: 1.4 Da; (5) Fragment ion tolerance: 0.8 Da; (6) SWISS-PROT Homo sapiens database (reviewed, downloaded in April 2014) was used.

2.7 Glycan Derivatization with Meladrazine: For derivatization of the glycans by Meladrazine [19, 20], the glycans were dissolved in 1 μ L H₂O and then 20 μ L Meladrazine (T3) labeling reagent (Mr 235.2115) was added (50 mM, methanol: acetic acid=3:1) at 60°C for 3 hours. For LC analysis, the samples were then dried and dissolved in 20 μ L 80% ACN for HILIC separation.

2.8 LC-MS Identification for Derivatized Glycans: LC-MS system used for glycan identification was the same as previously used for peptide identification but with regular HPLC column rather than

nano column. A HALO Penta-HILIC column (2.1 mm × 150 mm, 5 μm particle size) was used. Mobile phases A and B were 10 mM ammonium formate in water (pH 6.5) and 98% ACN containing 0.1% formic acid, respectively. Ten microliters of sample was injected at the initial gradient condition (80% mobile phase). The gradient was then ramped from 80 to 50% solvent B over 25 min with a total run time of 40 min. Glycans were eluted at 0.2 mL/min. The positive ion mode was used for analysis. The MS detection range was set at 800 to 2000 *m/z*. The ESI spray and capillary voltages were set at 5 kV and 35 V, respectively, with the stainless steel capillary heating to 275°C. The glycan structures were further confirmed by CID MS/MS using the LTQ Orbitrap Velos. Precursor ions were selected in data dependent acquisition (DDA) mode. The glycans were permethylated prior to MS/MS analysis for better ionization of the fragments and preventing glycan rearrangement in the gas phase. In order to isomerically separate permethylated glycans, a porous graphitic carbon (PGC) column (HyperCarb, 10 μm × 100 mm, 5 μm particle size) was utilized in LC-MS/MS analysis.

2.9 Data Analysis of Glycans: Theoretical structures and *m/z* values of derivatized glycans were generated by GlycoWorkbench Software developed by the EUROCarbDB. The experimental value was used to confirm the results. The identified glycans were quantified manually with Xcalibur Qual Browser 2.1 (Thermo Fisher Scientific, San Jose, CA) using the peak area from the extracted ion chromatogram (XIC) with the following settings: (1) peaks were extracted with a 1 Da (±0.5 Da) mass window, (2) scan filter was set as full MS, and (3) Genesis peak detection algorithm was used. Nomenclature of glycan is used according to Essentials of Glycobiology[21] and the abbreviations are used according to NIBRT GlycoBase.

3 Results and Discussion

There are two different types of fucosylation structures of N-linked glycoprotein, which have been of interest in biomarker studies, i.e. core fucosylation and antennary fucosylation. It is often difficult to distinguish the two structures using direct ESI-MS/MS due to the low fragmentation signals in MS/MS analysis for glycans. The glycosylation is also known to be site-specific[16, 22]. We thus sought to develop a method to analyze site specific fucosylation of a model protein A1AT (alpha-1-antitrypsin) to distinguish core fucosylation and antennary fucosylation structures at each glycosite. A work-flow describing the methods used is shown in **Figure 1**.

Sialidase releases α (2-3,6,8,9)-linked N-acetylneuraminic acid leaving galactose as the terminal of the N-glycan. Subsequently, β -galactosidase cleaves β (1-4,6)-linked galactose on condition that no fucose is bound to the subterminal N-acetylglucosamine in an N-glycan, thus providing a means to distinguish core fucosylation and antennary fucosylation (as illustrated in **Figure 2**). The elongation of polylactosamine-type structure is not considered in this study. A previous study about the glycan structures of A1AT has shown that no polylactosamine-type structure has been found in human serum[23]. In our analysis, no obvious polylactosamine-type structure was detected. A list of sialidase or sialidase/galactosidase digested N-glycans are shown in **Table 1** and the theoretical m/z is calculated using GlycoWorkbench. In order to further enhance the ionization efficiency of glycans for detection, a Meladrazine derivatization reagent was used. In the Thermo LTQ ion trap instrument, the signal in the m/z range of 1000-2000 is more intense than the lower or the higher mass ranges. The m/z within this range is indicated in bold. As shown in **Table 1** the core-fucosylated glycan and the antennary-fucosylated glycan have the same m/z in sialidase digested sample while the two have different m/z in the sialidase/galactosidase double digested sample. Therefore with sialidase/galactosidase double digestion the two types of fucosylation are readily distinguished without further MS/MS analysis or extensive sequential exoglycosidase digestion[24].

Further MS/MS analysis and fucosidase digestion analysis were also performed to confirm this result. The MS/MS data are shown in the supporting information. The MS/MS spectra of core-fucosylated and antennary-fucosylated bi-antennary and tri-antennary glycans are shown in the supporting information Figure S1. It could be clearly seen that galactosidase was not able to cut the galactose off if an antennary-fucosylation was formed, leaving both galactose and fucose on the glycan; whereas the fucosylation that formed at the core position does not have galactose left after galactosidase digestion. At the same time, this was also confirmed by fucosidase digestion analysis. The XIC of tri-antennary glycans, either core-fucosylated or antennary-fucosylated were extracted after various fucosidase digestions, as shown in the supporting information Figure S2. After core fucosidase α 1-2,3,4,6 fucosidase digestion (which cuts core fucosylation more efficiently than antennary fucosylation), the peak of core-fucosylated tri-antennary glycan(FA3) fully disappeared and some of the antennary-fucosylated tri-antennary glycan (A3FG) also disappeared with the appearance of its product glycan (A3G). After antennary fucosidase (α 1-2 fucosidase and α 1-3,4 fucosidase) digestion, the peak of FA3 did not change whereas most of A3FG disappeared with the appearance of A3G. This result confirmed the results from our method.

In this study A1AT was used as the model protein. The tri-antennary glycan of A1AT was used as an example as shown in **Figure 3**. Two aliquots of 10 μ g A1AT were digested with a single sialidase or with sialidase/galactosidase double digestion in parallel for comparison. The base peak of the theoretical m/z of a glycan was extracted manually using Xcalibur. With a single sialidase digestion, one peak with m/z 1203.49 (FA3G3 or A3FG3, the abbreviations of glycan structure are used according to NIBRT GlycoBase) was found, indicating there is one fucose attached to the tri-antennary glycan, but it is unknown whether it is core fucosylation or antennary fucosylation. With sialidase/galactosidase, the peak with m/z 1203.49 disappeared, and a minor peak with m/z 1901.83 (FA3) and a major peak with m/z 1041.45 (A3FG) were observed, indicating A1AT contains both a

core-fucosylated tri-antennary glycan at lower abundance and an antennary-fucosylated tri-antennary glycan at higher abundance. The completeness of galactosidase digestion is verified by extracting the XIC of the half-digested glycan structure with m/z at 1113.48 (FA3G2 or A3FG2). It can be seen clearly that there is no half-digested glycan structure so that galactosidase digestion is complete.

Analysis of A1AT showed that there are 12 different glycan structures after sialidase/galactosidase double digestion as shown in **Table 2**. The retention time of a core-fucosylated glycan is around 30 s longer than its corresponding non-fucosylated glycan; while the antennary-fucosylated glycan is around 100 s longer than its corresponding non-fucosylated glycan, indicating that the addition of one core-fucose enhanced the hydrophilicity of the glycopeptides and the addition of an antennary-fucose and a galactose enhanced its hydrophilicity further. The nonfucosylated biantennary and tri-antennary glycans are the top two most abundant glycan structures of A1AT, comprising 35.8% and 25.2% respectively. Both core and antennary fucosylation were found, with overall antennary fucosylation levels much higher than core fucosylation levels. In a classical lectin blot assay, one study has found that the up-regulation of core-fucosylated but not antennary-fucosylated A1AT could be indicative for hepatocellular cancer diagnosis [9]. The procedure developed in this study would enable the precise identification and quantification of core and antennary fucosylation of A1AT.

It is not known yet which site has core fucosylation or antennary fucosylation and the relative amount of the two fucosylations at each glycosite. A1AT has three theoretical N-glycosylation sites, all of which are detectable with trypsin digestion: site 70, site 107 and site 271. In order to determine the site specificity and the fucosylation type of each site, A1AT was first digested with trypsin, followed by sialidase/galactosidase double digestion. The A1AT glycopeptide mixture was then fractionated with a HILIC column into 30 fractions prior to PNGase F digestion. Each fraction was divided into two equal parts. One set of all the fractions was used for peptide identification, and the other set was used for further glycan identification. The peptide of each fraction was identified by LC-MS/MS and

was quantified by #PSMs (the number of peptide spectrum matches) from ProteomeDiscoverer1.2 search engine. With PNGase F digestion, previously glycosylated peptides are deamidated showing +1 Da dynamic modification at the site of asparagines whereas the nonglycosylated peptides do not have this modification. The aim of HILIC separation is to separate different peptides rather than a quantitative comparison of glycosylated and non-glycosylated peptides. Considering the chemical deamidation rate is less than 7%[25], where our experiment showed that the amount of deamidated peptides was about twice the amount of non-deamidated peptides, chemical deamidation was not specifically differentiated with glycosylated peptides. Most of these deamidated peptides in the fraction were from glycosylated peptides rather than chemical deamidation.

All fractions have been checked by LC-MS/MS. The result showed that the glycopeptides with truncated glycans were found in the same fraction as non-glycopeptides, where the fraction was collected at every 2 min (**Figure 4**). The non-glycosylated peptides may not elute at the same retention time as glycosylated peptides but the two eluted at similar retention time and in the same fraction with the Xbridge HILIC column in our experimental system. Therefore, the glycan structures of one site can be identified in sequential fractions. Three fractions F17 (fraction 17), F20 and F26 have maximum #PSMs of three glycopeptides of A1AT. As shown in **Table 3**, F17 mainly contains the glycopeptide YLGnATAIFFLPDEGK with site 271; F20 mainly contains the glycopeptide ADTHDEILEGLNFnLTEIPEAQIHEGFQELLR with site 107; F26 mainly contains the glycopeptide QLAHQSnSTNIFFSPVSIATAFAMLSLGTK with site 70. The MS/MS spectrum of one of the identified peptides of A1AT with glycosylation site 271, YLGnATAIFFLPDEGK (The deamidated asparagine is indicated by lower case) is shown in supporting information **Figure S3**. The glycan structures of the three fractions were further analyzed with LC-MS. The glycan was first extracted by PGC tips following PNGase F digestion and derivatized with T3 labeling reagent. Three glycan structures were identified in F17, including A2, FA2 and A3 (**Figure 5A**). Five glycan structures were identified in F20,

including A2, A3, A4, A3FG, A4FG (**Figure 5B**); No glycans were identified in F26, where within the limits of our current experiments no glycans were observed for site 70. The mass spectra of glycans are shown in supporting information **Figure S4**.

4 Concluding Remarks

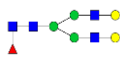
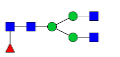

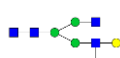






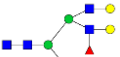
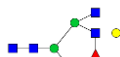




We have developed a work-flow to study the glycosylation of A1AT to identify the presence of core versus antennary fucosylation. This was performed on a standard A1AT protein which was digested by trypsin followed by sialidase/galactosidase digestion. Galactosidase removes terminal galactose residues in an N-glycan but does not cut when the subterminal N-acetylglucosamine is modified by fucosylation, thus providing a means to distinguish core fucosylation and antennary fucosylation. The resulting peptides were separated by HILIC separation and collected into 30 fractions. For each fraction the glycans were released using PNGase F digestion. The glycans could then be extracted and analyzed using Meladrazine T3 labeling reagent and HILIC-LC/MS where glycans with core or antennary fucosylation could be identified via mass spectrometry. In addition, the site of glycosylation could be studied via C18 extraction of the peptides followed by LC-MS/MS analysis. We identified 12 different glycans at two different glycosites in A1AT following sialidase/galactosidase digestion which clearly showed the presence of core versus antennary fucosylation. In future work, this methodology will be used to study changes in serum A1AT glycosylation during the progression of various cancers.

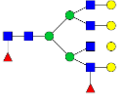
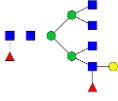
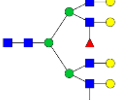
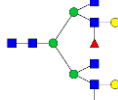
Acknowledgements: This work was funded by the National Cancer Institute under grant R01 CA160254 (DML) and received partial support from the National Cancer Institute under grant R01CA154455 (DML) and from the National Institutes of Health through grant R01 GM 49500 (DML) and grant 1R01GM112490-02(YM).

There are no conflicts of interest to declare.

Table 1. The list of sialidase or sialidase/galactosidase digested N-glycans and their theoretical m/z .

The m/z within 1000-2000 mass range is indicated in bold.

Glycan Type	Digest with Sialidase		Digest with Sialidase/Galactosidase	
	$[M-T3+H]^+$ or $[M-T3+2H+H_2O]^{2+}$	Structure	$[M-T3+H]^+$ or $[M-T3+2H+H_2O]^{2+}$	Structure
single fuco biantennary	2022.86/ 1020.93		1698.75 /858.88	
			1860.81 /939.91	
single fuco tri- antennary	2387.99/ 1203.49		1901.83 /960.42	
			2063.89/ 1041.45	
double fuco tri-antennary	2534.05/ 1276.52		2209.95/ 1114.48	
			2372.01/ 1195.51	
single fuco tetra- antennary	2753.12/ 1386.06		2104.93/ 1061.97	
			2266.97/ 1142.99	

double fuco tetra- antennary		2413.02/ <u>1216.01</u>	
	2899.18/ <u>1459.09</u>		
		2575.09/ <u>1297.05</u>	

blue square, GlcNAc; green circle, mannose; yellow circle, galactose; red triangle, fucose

Table 2. Summary of truncated N-glycans identified in A1AT, including *m/z*, retention time and % relative peak area.

Glycan Structure	[M-T3+H] ⁺ or [M-T3+2H+H ₂ O] ²⁺	Retention Time	% Relative Peak Area
A2	1552.7	7.72	35.8%
FA2	1698.8	8.37	1.8%
A2FG	1860.8	9.66	0.5%
A3	1755.8	8.43	25.2%
FA3	1901.8	8.99	0.7%
A3FG	1041.5	10.2	11.3%
A4	1958.8	9.24	7.0%
A4FG	1142.9	10.92	13.6%
A3F2G2	1195.5	11.74	0.4%
FA3FG	1114.5	10.69	0.1%
A4F2G2	1297.0	12.05	3.6%
FA4FG	1215.3	11.36	0.1%

Table 3. The peptide identified in fractions of HILIC column fractionation. The glycosylation site is indicated by lower case.

Fraction No.	Site	Peptide
F26	site 70	QLAHQSnSTNIFFSPVSIATAFAMLSLGTK
F20	site 107	ADTHDEILEGLNFnLTEIPEAQIHEGFQELLR
F17	site 271	YLGnATAIFFLPDEGK

Figure Captions

Figure 1. Work-flow of the experiment for determining glycosylation of A1AT. The glycosylated A1AT was first digested into peptides, followed by glycan truncation by sialidase/galactosidase double digestion. The glycopeptides were fractionated using a HILIC column. PNGase F digestion was used for separate mass spectral analysis of glycans and peptides.

Figure 2. Schematic of digestion by galactosidase which illustrates that digestion occurs for core fucosylation but not antennary fucosylation.

Figure 3. The fucosylation structure of fucosylated tri-antennary glycan treated by sialidase digestion (A) cannot be distinguished, whereas with sialidase/galactosidase double digestion (B) the core-fucosylated tri-antennary glycan with m/z at 1901.83 (FA3) elutes earlier than antennary-fucosylated tri-antennary glycan with m/z at 1041.45 (A3FG). The completeness of galactosidase digestion is verified by extracting the XIC of the half-digested glycan structure with m/z at 1113.48 (FA3G2 or A3FG2).

Figure 4. The #PSMs of identified glycosylated peptides or non-glycosylated peptides in fractions of HILIC fractionation. Deamidated peptides after PNGase digestion are considered glycosylated peptides, whereas non-glycosylated peptides are not deamidated. Chemical deamidated peptides were not differentiated due to their low percentile. “site70_deamidated” “site107_deamidated” and “site 271_deamidated” indicate glycosylated peptides with glycosylation site 70, 107 and 271 respectively “site 70” “site 107” and “site 271” indicate the corresponding non-glycosylated peptides.

Figure 5. LC-MS analysis of glycans in F17 (A) and F20 (B) from HILIC column fractionation. F17 contains A2, FA2 and A3; F20 contains A2, A3, A3FG, A4 and A4FG. No glycan structure was detected in F26.

Figure 1.

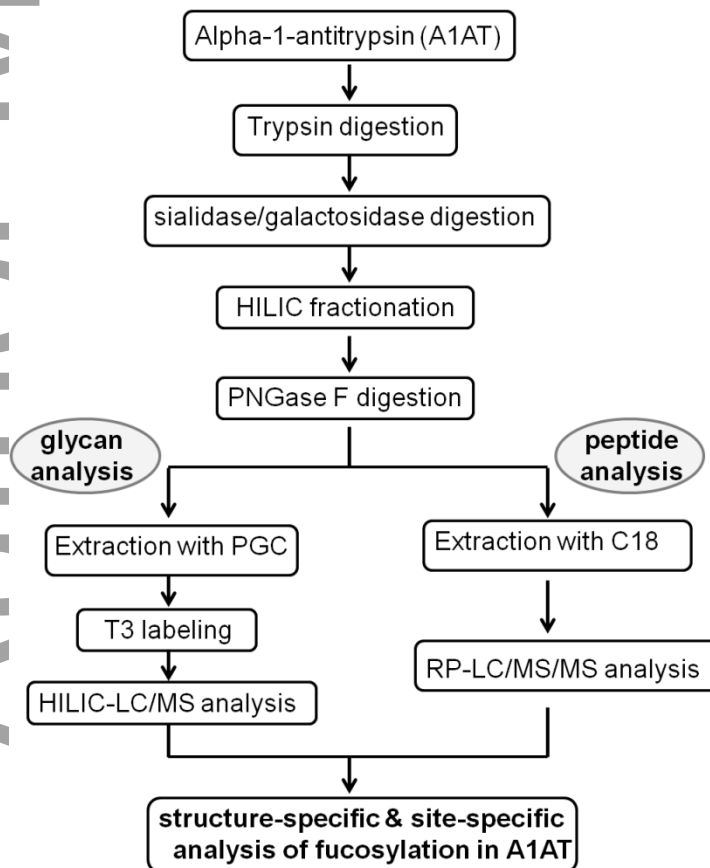


Figure 2.

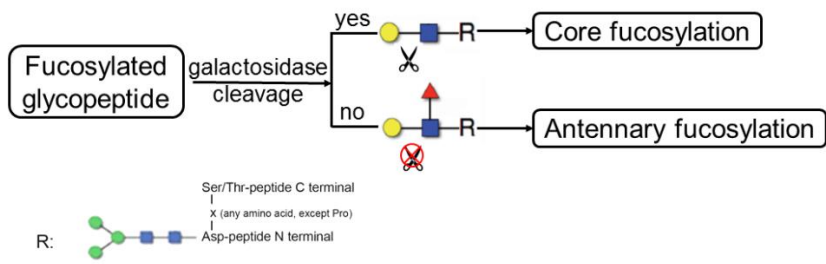
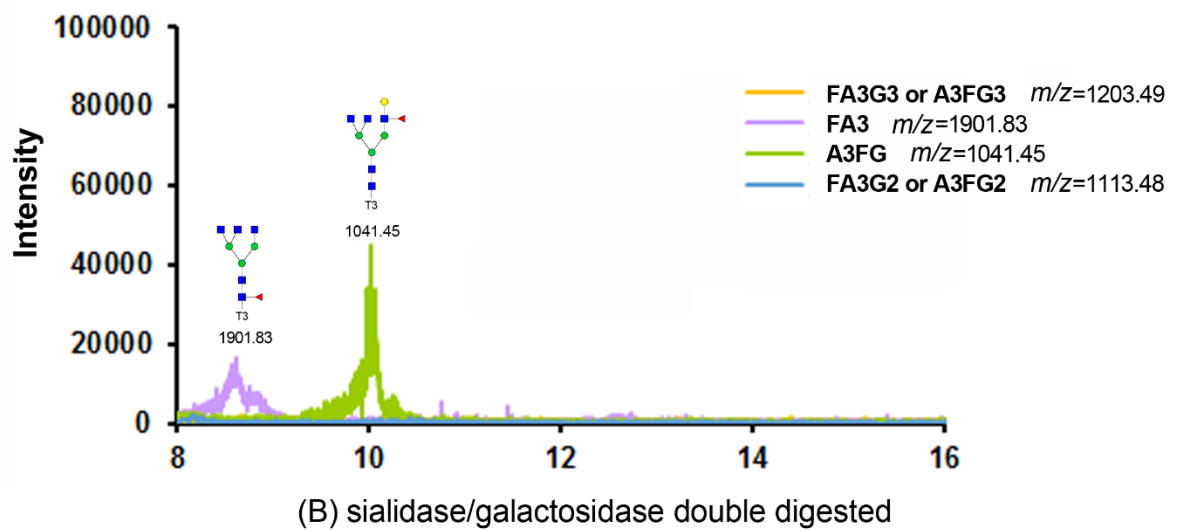
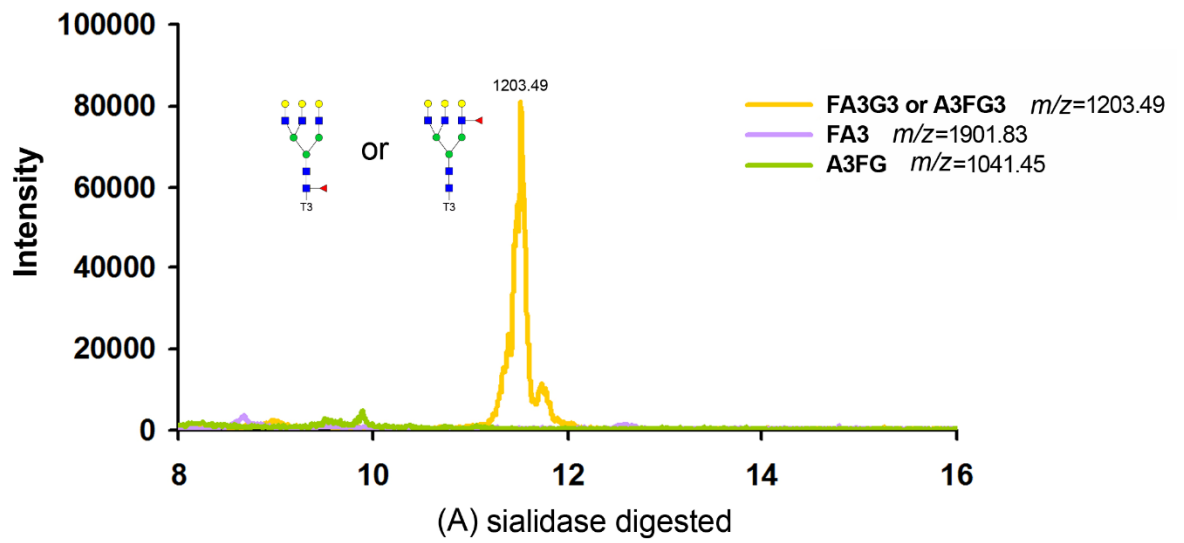
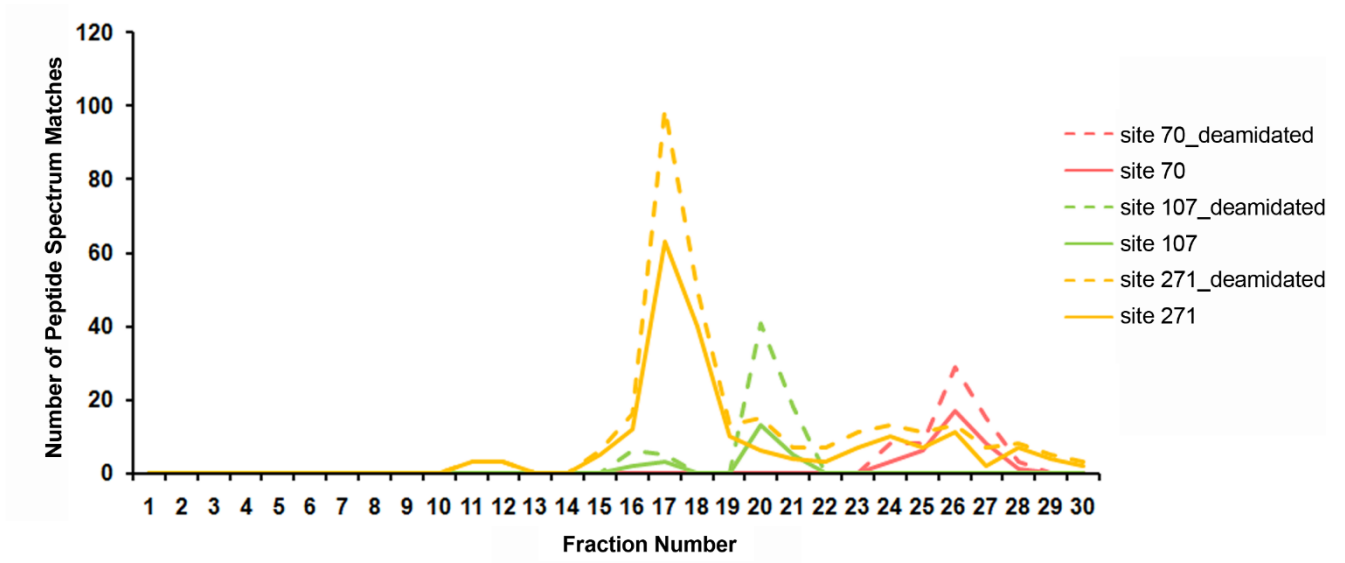


Figure 3.



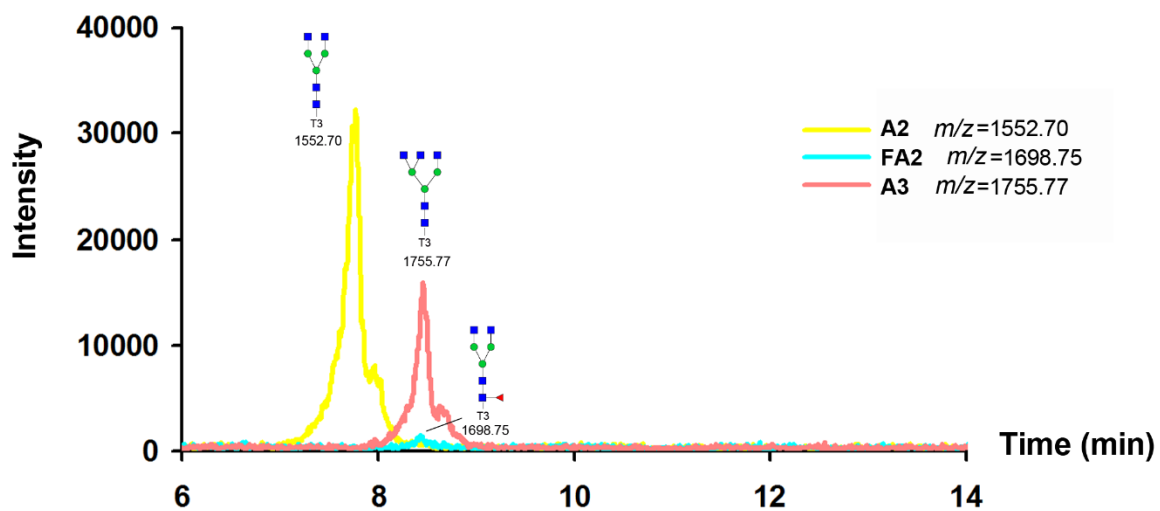
Auth

Figure 4.

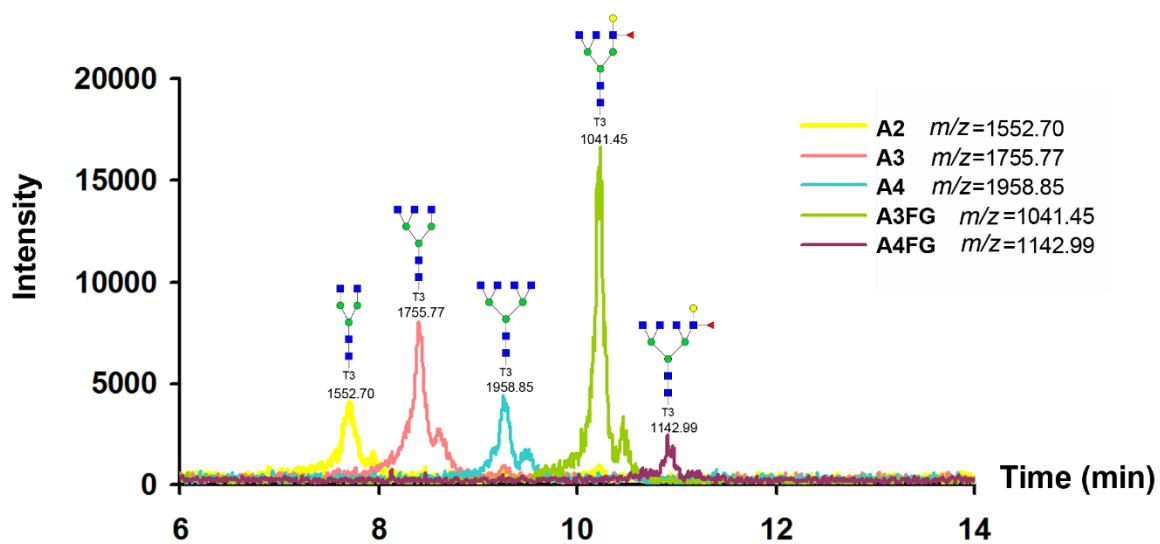


Author Man

Figure 5.



(A) F17



(B) F20

Auth

5 References

- [1] Lau, K. S., Dennis, J. W., *Glycobiology* 2008, *18*, 750-760.
- [2] Drake, R. R., *Advances in cancer research* 2015, *126*, 1-10.
- [3] Pompach, P., Brnakova, Z., Sanda, M., Wu, J., Edwards, N., Goldman, R., *Molecular & cellular proteomics : MCP* 2013, *12*, 1281-1293.
- [4] Okuyama, N., Ide, Y., Nakano, M., Nakagawa, T., Yamanaka, K., Moriwaki, K., Murata, K., Ohigashi, H., Yokoyama, S., Eguchi, H., Ishikawa, O., Ito, T., Kato, M., Kasahara, A., Kawano, S., Gu, J., Taniguchi, N., Miyoshi, E., *Int J Cancer* 2006, *118*, 2803-2808.
- [5] Liang, Y., Ma, T., Thakur, A., Yu, H., Gao, L., Shi, P., Li, X., Ren, H., Jia, L., Zhang, S., Li, Z., Chen, M., *Glycobiology* 2015, *25*, 331-340.
- [6] Jiang, K., Shang, S., Li, W., Guo, K., Qin, X., Zhang, S., Liu, Y., *Glycoconjugate journal* 2015, *32*, 657-664.
- [7] Haab, B. B., *Proteomics. Clinical applications* 2012, *6*, 346-350.
- [8] Wu, J., Xie, X., Liu, Y., He, J., Benitez, R., Buckanovich, R. J., Lubman, D. M., *J Proteome Res* 2012, *11*, 4541-4552.
- [9] Comunale, M. A., Rodemich-Betesh, L., Hafner, J., Wang, M., Norton, P., Di Bisceglie, A. M., Block, T., Mehta, A., *PLoS one* 2010, *5*, e12419.
- [10] Zhu, J., Lin, Z., Wu, J., Yin, H., Dai, J., Feng, Z., Marrero, J., Lubman, D. M., *J Proteome Res* 2014, *13*, 2986-2997.
- [11] Tian, Y., Zhang, H., *Proteomics. Clinical applications* 2010, *4*, 124-132.
- [12] Nie, S., Lo, A., Wu, J., Zhu, J., Tan, Z., Simeone, D. M., Anderson, M. A., Shedden, K. A., Ruffin, M. T., Lubman, D. M., *J Proteome Res* 2014, *13*, 1873-1884.
- [13] Froehlich, J. W., Barboza, M., Chu, C., Lerno, L. A., Jr., Clowers, B. H., Zivkovic, A. M., German, J. B., Lebrilla, C. B., *Anal Chem* 2011, *83*, 5541-5547.
- [14] Chandler, K. B., Pompach, P., Goldman, R., Edwards, N., *J Proteome Res* 2013, *12*, 3652-3666.
- [15] Wu, Y., Mechref, Y., Klouckova, I., Mayampurath, A., Novotny, M. V., Tang, H., *Rapid communications in mass spectrometry : RCM* 2010, *24*, 965-972.
- [16] Yin, H., Lin, Z., Nie, S., Wu, J., Tan, Z., Zhu, J., Dai, J., Feng, Z., Marrero, J., Lubman, D. M., *J Proteome Res* 2014, *13*, 2887-2896.
- [17] Yin, H., Tan, Z., Wu, J., Zhu, J., Shedden, K. A., Marrero, J., Lubman, D. M., *J Proteome Res* 2015, *14*, 4876-4884.
- [18] Varadi, C., Mittermayr, S., Szekrenyes, A., Kadas, J., Takacs, L., Kurucz, I., Guttman, A., *Electrophoresis* 2013, *34*, 2287-2294.
- [19] Zhang, Y., Zhu, J., Yin, H., Marrero, J., Zhang, X. X., Lubman, D. M., *J Proteome Res* 2015, *14*, 5388-5395.
- [20] Tie, C., Zhang, X.-X., *Analytical Methods* 2012, *4*, 357-359.

- [21] Varki, A., Cummings, D. R., Esko, D. J., Freeze, H. H., Stanley, P., Bertozzi, R. C., Hart, W. G., Etzler, E. M., *Essentials of Glycobiology*, Cold Spring Harbor Laboratory Press, New York 2008.
- [22] Pompach, P., Ashline, D. J., Brnakova, Z., Benicky, J., Sanda, M., Goldman, R., *J Proteome Res* 2014, *13*, 5561-5569.
- [23] McCarthy, C., Saldova, R., O'Brien, M. E., Bergin, D. A., Carroll, T. P., Keenan, J., Meleady, P., Henry, M., Clynes, M., Rudd, P. M., Reeves, E. P., McElvaney, N. G., *J Proteome Res* 2014, *13*, 596-605.
- [24] Guttman, A., *Electrophoresis* 1997, *18*, 1136-1141.
- [25] Zielinska, D. F., Gnad, F., Wisniewski, J. R., Mann, M., *Cell* 2010, *141*, 897-907.

Author Manuscript