# Naturally Occurring Structural Isomers in Serum IgA1 *O*-Glycosylation

**Kazuo Takahashi**[2,1], **Archer D. Smith IV**[1], **Knud Poulsen**[3], **Mogens Kilian**[3], **Bruce A. Julian**[4,2], **Jiri Mestecky**[2,4], **Jan Novak**[2], and **Matthew B. Renfrow**[1,*]

Matthew B. Renfrow: Renfrow@uab.edu

[1]Biomedical FT-ICR Mass Spectrometry Laboratory, Department of Biochemistry and Molecular Genetics, Aarhus University, Aarhus, Denmark

[2]Department of Microbiology, Aarhus University, Aarhus, Denmark

[3]Department of Medical Microbiology & Immunology, Aarhus University, Aarhus, Denmark

[4]Department of Medicine, Aarhus University, Aarhus, Denmark

## Abstract

IgA is the most abundantly produced antibody and plays an important role in the mucosal immune system. Human IgA is represented by two isotypes, IgA1 and IgA2. The major structural difference between these two subclasses is the presence of nine potential sites of *O*-glycosylation in the hinge region between the first and second constant region domains of the heavy chain. $Thr^{225}$, $Thr^{228}$, $Ser^{230}$, $Ser^{232}$ and $Thr^{236}$ have been identified as the predominant sites of *O*-glycan attachment. The range and distribution of *O*-glycan chains at each site within the context of adjacent sites in this clustered region create a complex heterogeneity of surface epitopes that is incompletely defined. We previously described the analysis of IgA1 *O*-glycan heterogeneity by use of high resolution LC/MS and electron capture dissociation tandem MS to unambiguously localize all amino acid attachment sites in IgA1 (Ale) myeloma protein. Here, we report the identification and elucidation of IgA1 *O*-glycopeptide structural isomers that occur based on amino acid position of the attached glycans (positional isomers) and the structure of the *O*-glycan chains at individual sites (glycan isomers). These isomers are present in a model IgA1 (Mce1) myeloma protein and occur naturally in normal human serum IgA1. Variable *O*-glycan chains attached to $Ser^{230}$, $Thr^{233}$ or $Thr^{236}$ produce the predominant positional isomers, including *O*-glycans composed of a single GalNAc residue. These findings represent the first definitive identification of structural isomeric IgA1 *O*-glycoforms, define the single-site heterogeneity for all *O*-glycan sites in a single sample, and have implications for defining epitopes based on clustered *O*-glycan variability.

## Keywords

*O*-glycosylation; IgA1 immunoglobulin; ECD; electron capture dissociation; FT-ICR MS; glycopeptide isomers

---

*To whom correspondence should be addressed. Matthew B. Renfrow, PhD, UAB Biomedical FT-ICR MS Laboratory, MCLM 570, 1530 3rd Avenue South, Birmingham, AL 35294. Phone: 205-996-4681, Fax: 205-975-2188; renfrow@uab.edu. or Jan Novak, PhD, 845 19th Street South, BBRB 734, Birmingham, AL 35294. Phone: 205-934-4480, Fax: 205-934-3894; jannovak@uab.edu.

## Introduction

IgA is the most abundant isotype of human immunoglobulins and has a significant role in mucosal immunity. Approximately two thirds of IgA are produced in a polymeric form in mucosal tissues, particularly in the intestinal tract, and then selectively transported by a receptor-mediated pathway into external secretions [1]. In contrast, most of circulatory IgA occurs in a monomeric form. Production of IgA in the mucosal compartments is a dominant immunological process important for homeostasis between the gut commensal microbiota and the local immunological environment [2]. Specifically, IgA can neutralize toxins and pathogenic microbes as well as restrict the gut commensal flora to the intestinal lumen.

Human IgA is represented by two structurally and functionally distinct subclasses, IgA1 and IgA2 [3]. The most significant difference between the two isotypes of IgA is that IgA1, but not IgA2, possesses a 19-amin*O*-acid hinge region (HR) with 9 potential *O*-glycosylation sites, of which 3 to 5 are usually occupied by core 1 type *O*-glycans [4, 5] (Fig. 1A). Mattu et al. determined that *O*-glycans in circulatory IgA1 are commonly attached at $Thr^{228}$, $Ser^{230}$, and $Ser^{232}$, whereas sites $Thr^{225}$ and $Thr^{236}$ are occupied in only a fraction of IgA1 molecules [5]. We previously reported on the direct localization of IgA1 (Ale) myeloma protein *O*-glycan chains to these same sites by use of activated ion-electron capture dissociation (AI-ECD) fragmentation. Additionally, several groups, including ours, have reported on the profile or distribution of IgA1 *O*-glycoforms present in a variety of IgA1 myeloma proteins [6-12], normal human serum IgA1 [13-17], and IgA1 from patients with IgAN [14, 18-20]. While some IgA1 myeloma proteins exhibit fewer *O*-glycans, the majority of these analyses indicate the native distribution of IgA1 HR *O*-glycoforms centers around four to five *O*-glycan chains with a total range of 3-6 *O*-glycans.

The IgA1 HR *O*-glycosylation is mediated by specific glycosyltransferases in the Golgi apparatus of IgA1-secreting cells. Iwasaki et al. reported that *O*-glycosylation of IgA1 is initiated by a UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase 2 (GalNAc-T2). Isomeric structures of synthetic IgA1 HR glycopeptides associated with sites of different GalNAc attachment were generated by recombinant GalNAc-T2 [21], whereas other GalNAc-T family members have been shown to add GalNAc to IgA1 HR as well[22]. GalNAc-T2 rarely glycosylates sites adjacent to already occupied sites in HR, such as $Ser^{232}$ and $Thr^{233}$. Interestingly, five other GalNAc-Ts that can add GalNAc at $Thr^{233}$ are expressed in IgA-positive cells [21]. Positions of *O*-glycans in glycoproteins are dependent on the variety of GalNAc-Ts expressed in the cells and their substrate specificities [23]. Differential regulation of these enzymes in IgA1-producing cell may be implicated in the generation of isomeric structures in IgA1 HR. The question remains as to whether attachment of all IgA1 *O*-glycan chains is mediated by the same GalNAc-T, *i.e.*, GalNAc-T2.

Regardless of which specific GalNAc-T(s) (in the Golgi apparatus) mediate the initial glycan attachment, some sites on IgA1 secreted by the IgA1-producing cells remain galactose (Gal) deficient [24, 25]. The levels of serum IgA1 with this aberration are increased in patients with IgAN [26-28] compared to those in normal human serum IgA1 [29]. IgAN is characterized by predominant IgA1 deposition in the renal mesangium [19, 30]. In patients with IgAN, IgA1 glycoproteins with Gal-deficient *O*-glycans in the HR are present in the circulating immune complexes and in mesangial immunodeposits. This Gal-deficient IgA1 is produced by IgA1-secreting cells [25], Although elevated levels of aberrantly glycosylated IgA1 in the sera of patients with IgAN have been reported [18-20, 27-29, 31-34], there have not been any systematic MS studies that have included site-specific heterogeneity information. To characterize the pathogenic forms of IgA1 involved with IgAN, *O*-glycan microheterogeneity and attachment sites including isomeric structures should be analyzed, as each HR has nine potential *O*-glycosylation sites. The possibility of structural isomers of

IgA1 $O$-glycosylation has potential impact in other areas as well including celiac disease [35] and the use of IgA as a therapeutic antibody [36, 37].

In our original analysis of IgA1 HR glycopeptides by use of AI-ECD FT-ICR tandem MS, we demonstrated the ability of ECD fragmentation to delineate between homogeneous glycopeptides ion populations and positional isomeric mixtures of a synthetic IgA1 HR glycopeptides with a single GalNAc residue at Thr$^{228}$ or Ser$^{230}$ [9, 38-40]. Still, isomeric structures of native IgA1 HR $O$-glycopeptides have not been well elucidated. Two types of isomeric structures can occur in clustered $O$-glycopeptides: isomers based on glycosidic-bond differences and amino acid positional isomers. Here, we demonstrate the first definitive identification of a series of naturally occurring structural isomers of HR $O$-glycopeptides in IgA1 myeloma protein and normal serum IgA1. Furthermore, we successfully identified mixed isomeric glycoforms associated with $N$-acetyl neuraminic acid (NeuAc) attachment to GalNAc or Gal. Our results demonstrated a defined relative quantitative distribution of isomeric IgA1 $O$-glycoforms based on alternative attachment of $O$-glycans to Ser$^{230}$, Thr$^{233}$ and Thr$^{236}$ in normal human serum. These results provide insight about the glycosylation pathways of human IgA1 $O$-glycans, have implications for studies of diseases with aberrant glycosylation, and define an unbiased process for defining clustered $O$-glycan heterogeneity at the monosacharride and disaccharide level.

## Materials and Methods

### Isolation of IgA1 and preparation of proteolytic fragments

Polymeric IgA1 (Mce1) myeloma protein was previously isolated from plasma of a patient with multiple myeloma using standard purification protocols [41]. Normal human serum IgA was purified using affinity chromatography with anti-human IgA (Cappel Laboratories, Cochranville, PA) coupled to CNBr-Sepharose column [42]. Purified samples were stored aliquoted at -80°C. Purified IgA1 proteins were treated with an IgA-specific protease (from *Clostridium ramosum* strain AK183, recombinant in *Escherichia coli; Streptococcus pneumoniae* strain TIGR4, recombinant in *E. coli*; or *Haemophilus influenzae*, strain HK50 that vary in the cleavage site; see Fig. 1A) followed by trypsin, to release IgA1 HR glycopeptides [12]. The digestion with each IgA-specific protease was performed for 24 h at 37°C. The digests of IgA1 (Mce1) myeloma protein were treated with 0.25 U/ml neuraminidase (ProZyme, Hayward, CA) in 50 mM sodium phosphate, pH 6.0, to remove sialic acid residues from the $O$-glycan chains [28]. After neuraminidase treatment, the preparations were reduced with 5 mM DTT. Then, trypsin (Promega, Madison, WI) was added for 18 h at 37°C in 100 mM NH$_4$HCO$_3$, pH 8.3. The digests were desalted by use of a C18 spin column (Pierce, Rockford, IL) before MS analyses. This study was approved by the Institutional Review Board of the University of Alabama at Birmingham.

### Offline LC separation

Twenty-five μg of digested IgA1 was fractionated by use of an offline reversed-phase liquid chromatography 300 μm × 15 cm C18 PepMap column (offline LC) (LC Packings, Sunnyvale, CA), as previously described [9, 43]. The digests were eluted with an acetonitrile gradient from 5% to 30% for IgA1 myeloma protein treated with HK50+trypsin, and 5% to 18% for IgA1 myeloma protein digested with TIGR4+trypsin and IgA1 purified from normal human serum, in 0.1% formic acid over 50 minutes at 5 μl min$^{-1}$. Eluted IgA1 HR glycopeptides were collected as 5-μL fractions into a 96-well plate by use of an Agilent 1100 micro-well plate fraction collector (Agilent 1100, Santa Clara, CA).

### Online LC FT-ICR MS analyses

On-line LC was performed by use of an Eksigent MicroAS autosampler and 2D LC nanopump (Eksigent, Dublin, CA). Three-hundred ng of digested IgA1 was loaded onto a 100-μm diameter, 11-cm column pulled tip packed with Jupiter 5-μm C18 reversed phase beads (Phenomenex, Torrance, CA). The digests were then eluted with an acetonitrile gradient from 5% to 30% in 0.1% formic acid over 50 minutes at 650 nl min$^{-1}$. Linear quadrupole ion trap (LTQ) FT-ICR (LTQ FT, Thermo Fisher Scientific, San Jose, CA) parameters were as described previously [10].

### AI-ECD FT-ICR tandem MS

Offline fractionated IgA1 HR glycopeptides were analyzed by use of a 7 T LTQ FT-ICR MS. A monolithic silicon microchip-based electrospray interface, the TriVersa ™ NanoMate (Advion, Ithaca, NY), served as the source for electrospray ionization (ESI) [10], and AI-ECD FT-ICR tandem MS analysis was performed as previously described with some modifications. Isolation width was 10 *Th* with an AGC target value of $2 \times 10^6$, maximum fill 3,500 ms. Following transfer to the ICR cell, precursor ion populations were photon-irradiated for 100 ms at 10% (2 W) laser power. After 50 ms of photon-irradiation, the precursor populations were irradiated with the electrons for 100 ms at 2-3% energy ($\approx 0.5$eV). Each AI-ECD scan was acquired as an FT-ICR broadband mass spectrum ($100 < m/z < 3,000$) at a mass resolving power of 100,000 at $m/z$ 400. Each displayed spectrum represents a sum of 100 scans.

### Data analysis

All spectra were analyzed by use of the Xcalibur Qual Browser 2.0 (Thermo Fisher Scientific) software. Individual IgA1 *O*-glycopeptides were identified as previously described [10, 12]. Briefly, IgA1 *O*-glycopeptides species in each LC/MS analysis were identified from the known sequence of the isolated HR glycopeptides, calculated monoisotopic mass of glycopeptides and the presence of adjacent HR glycopeptides within the high-resolution FT-ICR MS spectra. Monoisotopic $m/z$ values for the IgA1 *O*-glycopeptides were manually tabulated from the raw data files by use of Xcalibur Qual Browser 2.0. The minimal threshold for IgA1 HR glycopeptide identification and relative abundance measurements was a signal-to-noise ratio of at least 5:1 with >4 isotopic peaks. Identified IgA1 HR *O*-glycopeptides were checked against theoretical values by use of the GlycoMod tool (http://www.expasy.org). Known IgA1 HR amino acid sequences based on the combination of IgA-specific protease + trypsin digestions performed were inputted with trypsin enzyme and 0 missed cleavage sites. Hexose, *N*-acetylhexosamine (HexNAc) and NeuAc monosaccharide residues were all selected as possible (variable) additions to the IgA HR peptides with a mass tolerance of 10 ppm.

Relative quantitative distributions of each glycopeptide series were calculated by use of a label-free method previously described[12, 44]. After assigning all glycopeptides peaks in the spectrum, the ion chromatogram for each glycopeptides ion was individually extracted for the specific glycopeptide ion species $\pm$ 0.83 *Th*. Then the area under the peak for each glycopeptide's extracted ion chromatogram (XIC) was divided by the XIC for all glycopeptides of a given charge state. The values are reported as the percentage of each glycopeptide intensity relative to the total glycopeptide intensity as previously reported [12]. Each analysis of IgA-specific protease + trypsin preparations of IgA1 HR to obtain relative distribution values of IgA1 HR *O*-glycoforms was performed in triplicate. LC/MS replicate runs and differential digestions served as internal controls for assignments of relative abundance. Relative abundance of glycopeptides was expressed as the mean $\pm$ standard deviation (S.D.) of triplicate LC/MS runs. Theoretical lists of fragmented IgA1 HR peptides by AI-ECD were generated by use of the ProteinProspector MS product tool (http://

prospector.ucsf.edu/) with the inclusion of *c*, *z*, *b* and *y* ions. Peptide ion fragments of all IgA1 HR AI-ECD FT-ICR tandem MS spectra were manually assigned to locate sites of *O*-glycosylation and are provided in supplemental data.

## Results

### Identification of IgA1 O-glycopeptide amino acid positional isomers

Our standard analysis of IgA1 *O*-glycan heterogeneity involves two steps: 1) IgA1 HR *O*-glycopeptide high resolution LC/MS profiling to identify the number of unique IgA1 *O*-glycoforms within a single sample[10] and 2) characterization of individual IgA1 *O*-glycoforms in terms of sites of attachment and the composition of the *O*-glycan chain at each individual sites by use of AI-ECD FT-ICR tandem MS [9]. For the analysis of a model IgA1 (Mce1) myeloma protein and serum IgA1 isolated from two normal healthy volunteers, purified IgA1 was digested with three bacterial IgA-specific proteases, one at a time, followed by neuraminidase and trypsin. The digests, with the released series of HR *O*-glycopeptides, were then subjected to online LC FT-ICR MS. AK183, TIGR4 and HK50 (+ trypsin) generated N-terminal HR fragments ($His^{208}$-$Pro^{221}$, $His^{208}$-$Pro^{227}$, and $His^{208}$-$Pro^{231}$) and C-terminal HR fragments ($Val^{222}$-$Arg^{245}$, $Thr^{228}$-$Arg^{245}$, and $Ser^{232}$-$Arg^{245}$), respectively (Fig. 1A) as in our previous report [12].

For the HK50 + trypsin proteolytic fragment of IgA1 (Mce1) myeloma protein, nine glycopeptides corresponding to 1, 2 or 3 *O*-glycan chains attached to the proteolytic fragment $Ser^{232}$-$Arg^{245}$ (Fig. 1B) and five glycopeptides corresponding to 2 or 3 *O*-glycan chains attached to the $His^{208}$-$Arg^{245}$ proteolytic fragment (Supplemental Fig. 1A) were detected. The relative abundance of each IgA1 *O*-glycopeptide was calculated based on the area under the peak in the LC/MS XIC of individual glycopeptides. Each relative abundance is reported as a percentage against the total LC/MS XIC of all glycopeptides with the same amin-acid sequence within a single LC/MS analysis (three independent runs for each IgA protease + trypsin preparation) [12]. This method of label-free quantitative analysis of *O*-glycopeptides originating from the same proteolytic fragment provides a reproducible assessment of the distribution of *O*-glycoforms present and has been demonstrated for *N*-glycopeptides [44] as well as our previous work [12]. Table 1 provides a complete list of all observed $Ser^{232}$-$Arg^{245}$ IgA1 (Mce1) myeloma *O*-glycopeptides with their respective retention times, relative quantitative distributions, and mass errors based on three independent LC/MS runs. Based on these calculations, approximately 90% of the IgA1 HR in the sample was composed of glycoforms with $GalNAc_1Gal_1$, $GalNAc_2Gal_1$ or $GalNAc_2Gal_2$ residues attached to the C-terminal fragment ($Ser^{232}$-$Arg^{245}$). The remaining 10% of C-terminal glycopeptides included three glycosylated species with three GalNAc residues with one to three Gal residues. In the course of the label-free quantitative analysis, bimodal separation (two resolved peaks) of the XIC was observed for two IgA1 *O*-glycopeptides (*m/z* 678.961 and 732.979) corresponding to the attachment of $GalNAc_2Gal_1$ and $GalNAc_2Gal_2$ to the $Ser^{232}$-$Arg^{245}$ HR peptide, respectively (Fig. 1C). This was in contrast to a single peak observed in the XIC of the $Ser^{232}$-$Arg^{245}$ + $GalNAc_1Gal_1$ C-terminal HR peptide (*m/z* 611.267).

To characterize *O*-glycoforms with bimodal retention times in the online LC/MS analysis (Fig. 1C), the $Ser^{232}$-$Arg^{245}$ glycopeptides were fractionated by use of offline LC coupled to a microscale fraction collector [10]. Glycopeptides eluting at 28-30 min and 30-32 min were collected in adjacent fractions and then individually analyzed by direct infusion FT-ICR MS analysis (NanoMate™ ESI FT-ICR MS spectra are available in supplemental Fig. 2). The $Ser^{232}$-$Arg^{245}$ glycopeptides corresponding to the attachment of $GalNAc_1Gal_1$, $GalNAc_2Gal_1$, and $GalNAc_2Gal_2$ glycan residues were detected in both fractions. Each fraction was re-analyzed by online LC/MS separately and as an equal-volume mixture (1:1

v/v) under the same conditions as before (Fig. 1D). The XIC of $Ser^{232}$-$Arg^{245}$ + $GalNAc_1Gal_1$ glycopeptides ($m/z$ 611.267) showed a single peak in both fractions as well as in the equal-volume mixture, indicating the glycoform had identical hydrophobic characteristics in each microfraction. In contrast, the $Ser^{232}$-$Arg^{245}$ + $GalNAc_2Gal_1$ glycopeptide ($m/z$ 678.961) showed a single peak in the XIC for each adjacent fraction, but each had distinct retention times (XIC maximum ion abundance for each     30 seconds apart). When the two fractions were combined, a bimodal peak was clearly observed as seen in the unfractionated sample. A similar distribution was observed in the XIC of the $Ser^{232}$-$Arg^{245}$ + $GalNAc_2Gal_2$ glycopeptide ($m/z$ 732.979) with a bimodal peak in the earlier fraction (28-30 min) and a single peak in latter fraction (30-32 min). These results together indicated that offline LC fractionated glycopeptides included mixtures of isomeric glycoforms, *i.e.*, glycopeptides with the same composition but differing only in some of the glycan-attachment sites (Fig. 1C and D).

### AI-ECD tandem MS of isomeric IgA1 O-glycopeptides

To further characterize the three $Ser^{232}$-$Arg^{245}$ HR glycopeptides, the HR glycopeptide ion species in each offline LC fraction were individually subjected to AI-ECD FT-ICR tandem MS to localize the sites of *O*-glycan attachment. The assigned IgA1 HR *O*-glycopeptides would assume that the three glycoforms observed in the offline fractions have a single GalNAc-Gal disaccharide ($GalNAc_1Gal_1$), a disaccharide plus a GalNAc monosaccharide ($GalNAc_2Gal_1$), and then two GalNAc-Gal disaccharides ($GalNAc_2Gal_2$) attached.

AI-ECD fragmentation analysis of the [$Ser^{232}$-$Arg^{245}$ + $GalNAc_2Gal_1$]$^{3+}$ ions ($m/z$ 678.961) from the adjacent offline fractions is shown in Fig. 1E and F. For this HR *O*-glycopeptide, the first observed N-terminal fragment ion in each fraction is $c_6$ ($m/z$ 1166.533 in 28-30 min fraction, 1166.535 in 30-32 min fraction) that assigns a disaccharide and a monosaccharide N-terminal to $Pro^{237}$ ($Ser^{232}$, $Thr^{233}$, or $Thr^{236}$). From the C-terminus, the observed z fragments differ between the two offline LC microfractions even though the precursor ion $m/z$ is identical. Specifically, the $z+1_{10}$ fragment ion is different. In the 28-30 min fraction, the $z+1_{10}$ ion ($m/z$ 1272.537) corresponds to the mass of $Thr^{236}$-$Arg^{245}$ + GalNAc followed by a $z+1_{13}$ ion ($m/z$ 1567.690) with no additional sugar residues attached, therefore assigning a monosaccharide at $Thr^{236}$ and a disaccharide at $Ser^{232}$. In the 30-32 min fraction, the observed $z+1_{10}$ ion ($m/z$ 1069.457) is lower, corresponding to the mass of $Thr^{236}$-$Arg^{245}$ alone. This is followed by a $z+1_{13}$ ion (m/z 1567.688) corresponding to the mass of $Thr^{233}$-$Arg^{245}$ + GalNAc, assigning the monosaccharide *O*-glycan to $Thr^{233}$. Fig. 1F shows the presence or absence of the distinguishing $z+1_{10}$ fragments from each AI-ECD FT-ICR MS spectrum. The combined results clearly identify $Ser^{232}$-$Arg^{245}$ IgA1 HR *O*-glycopeptide isomers based on the *O*-glycan amino acid position.in agreement with the bimodal XIC peaks. Subsequent mixing of the fractions at different ratios confirmed the relative abundance of the two isomeric *O*-glycoforms estimated from XIC. The observed AI-ECD fragments from each of the $Ser^{232}$-$Arg^{245}$ IgA1 HR *O*-glycopeptides (isomers and non-isomers) are detailed on the amin*O*-acid sequence of $Ser^{232}$-$Arg^{245}$ in supplemental Fig. 3. The full AI-ECD FT-ICR tandem MS spectra are provided in supplemental data.

Following the same methodology, we also performed AI-ECD tandem MS analysis on a larger series glycopeptides ($Thr^{228}$-$Arg^{245}$) generated by TIGR4 IgA-specific protease and trypsin to create an overlapping relative distribution as we have previously described[12]. The overlapping IgA HR peptides serve to corroborate assigned relative distributions and provides the context of two additional sites of *O*-glycosylation at $Thr^{228}$ and $Ser^{230}$. The *O*-glycan profile of $Thr^{228}$-$Arg^{245}$ HR was analyzed by online LC FT-ICR MS (Supplemental Table 2) and indicated that approximately 90% of glycopeptides were $GalNAc_3Gal_2$, $GalNAc_3Gal_3$, $GalNAc_4Gal_2$, $GalNAc_4Gal_3$ and $GalNAc_4Gal_4$. The remaining glycosylated species were those with five GalNAc residues and two to five Gal residues. The XIC of the

five major Thr$^{228}$-Arg$^{245}$ glycoforms showed multimodal distributions in all but the GalNAc$_3$Gal$_2$ glycoform (Fig. 2). *O*-glycan amino acid positional isomers were identified between Thr$^{233}$ and Thr$^{236}$ as before and a third site at Ser$^{230}$. Five molecular mass values corresponded to 11 distinct glycoforms. The relative distribution of all observed Thr$^{228}$-Arg$^{245}$ HR *O*-glycoforms are provided in Supplemental Table 2 as well as all AI-ECD FT-ICR MS/MS spectra in the supplemental data that assigned the sites of *O*-glycan attachment for all amino acid positional isomers. The relative quantitative distributions of a third corroborating series IgA1 proteolytic fragments (AK183 + trypsin release) are provided in Supplemental Table 3.

## Isomeric glycoforms of HR O-glycopeptides from normal human serum IgA1

Based on the discovery of these sets of isomeric IgA1 *O*-glycopeptides, we were interested in finding out if these isomers were unique to this myeloma protein or in fact naturally occurring in normal human serum IgA1. IgA1 was isolated from the serum of two healthy volunteers with no history of kidney disorder. Each sample was prepared and analyzed separately as before with the exception that neither sample was treated with neuraminidase. Thus in the context of identifying *O*-glycan isomers we hypothesized that a Ser$^{232}$-Arg$^{245}$ +GalNAc$_1$Gal$_1$NeuAc$_1$ trisaccharide could have isomers at the amino acid level and the glycosidic bond level. Normal human IgA1 digested with HK50+typsin (without neuraminidase treatment) were subjected to online LC FT-ICR MS and offline LC plus AI-ECD FT-ICR tandem MS as before. Perdivara *et al* demonstrated the successful localization of a single sialic acid residue by use of ECD fragmentation [45]. Thus we attempted to do the same to begin the process of defining the site-specific sialic acid heterogeneity in IgA1 HR clustered *O*-glycans while also determining if amino acid positional isomers occurred in normal human serum IgA1. Table 2 lists the results for the Ser$^{232}$-Arg$^{245}$ proteolytic fragment. The same series of isomers were found along with unambiguous localization of a sialic acid residue to Ser$^{232}$. Twenty-two glycopeptides with 1, 2 or 3 *O*-glycans in Ser$^{232}$-Arg$^{245}$ and twelve glycopeptides with 2 or 3 *O*-glycans in His$^{208}$-Arg$^{245}$ were detected (Table 2. LC/MS spectra are shown in Supplemental Fig. 4). The XIC of four ion species displayed the same bimodal distribution as seen before including two ion species with monosialylation (Fig. 3). Subsequent AI-ECD tandem MS analysis of offline LC fractions identified the same pairs of Thr$^{233}$ and Thr$^{236}$ isomers in both samples as well as unambiguously localizing the NeuAc residue to Ser$^{232}$. The AI-ECD assigned *O*-glycans are listed in Table 2 for one of the two IgA1 samples from serum of normal healthy controls. All AI-ECD tandem MS spectra are provided in supplemental data. The complete list of Ser$^{232}$-Arg$^{245}$ IgA1 HR *O*-glycoforms identified including di- and trisialylated species are provided in Supplemental Table 4. The relative distributions for both normal healthy serum IgA1 samples from serum of normal healthy controls are provided in Supplemental Table 5.

## Isomeric glycoforms associated with the site of NeuAc attachment

IRMPD tandem MS cleaves mainly glycosidic bonds, thereby allowing structural characterization of glycans [46, 47]. To identify isomeric structures based on the site of NeuAc attachment to Gal or GalNAc on normal serum IgA1, monosialylated glycoforms isolated in offline LC fractions were individually subjected to IRMPD FT-ICR tandem MS. Primary loss of NeuAc (291.095 Da) or Gal (162.053 Da) from the gas-phase isolated precursor ion indicated the *O*-glycans had two terminal sugars and thus were a mixture of isomeric structures having NeuAc attached in a linear trisaccharide (GalNAc-Gal-NeuAc) or a branched trisaccharide (GalNAc (NeuAc) -Gal) in the GalNAc$_1$Gal$_1$NeuAc$_1$ Ser$^{232}$-Arg$^{245}$ glycoform (Fig. 4). Evidence for the same NeuAc isomers can be seen in the IRMPD tandem MS spectra of the GalNAc$_2$Gal$_1$NeuAc$_1$ Ser$^{232}$-Arg$^{245}$ glycoform (Supplmental Fig. 5A). An additional GalNAc$_2$Gal$_2$NeuAc$_1$ glycoform having either fragment ion of

GalNAc$_2$Gal$_2$+HR or GalNAc$_2$NeuAc$_1$+HR also indicated the mixture of isomeric structures based on sites of NeuAc attachment (Supplemental Fig. 5B).

### Comparing IgA(Mce1) myeloma protein and normal human serum IgA1 amino acid positional isomers

After clearly identifying amino acid positional isomers in both IgA1 myeloma protein and IgA1 from sera of normal healthy controls, we wanted to compare the distributions of each population of IgA1 *O*-glycoforms. However, each sample was prepared differently (with and without neuraminidase). To remedy this discrepancy, we performed online LC/MS of fully sialylated IgA (Mce1) myeloma protein for the Val$^{222}$-Arg$^{245}$ fragment that encompasses all six identified sites of *O*-glycosylation. Thirty-nine distinct ion species were detected with 4-6 glycan chains attached including as many as four NeuAc residues attached (Supplemental Table 6). When the relative distriburions were assigned to all *O*-glycoforms, an interesting trend was observed. The identified IgA1 HR *O*-glycoforms were grouped based on their identical number of GalNAc and Gal residues. The combined percentage distributions (simply added together) closely mimicked the distributions observed for the desialylated profiles. While this makes sense from a population analysis perspective, it was empirically surprising due to the presence of 1-4 negatively charged NeuAc residues. This result suggested that the sialic acid residues have limited affect on the MS response of these heavily glycosylated peptides and that the MS response is instead dominated by the the HR protonated amino acid residues (3+ and 4+ charged series). This trend was observed for each of our IgA1 HR proteolytic preparations. Based on this analysis we compared the distribution of single-site *O*-glycan heterogeneity between IgA1 (Mce1) myeloma protein and normal human serum IgA1 at Ser$^{232}$, Thr$^{233}$ and Thr$^{236}$. Fig. 5 clearly shows that a disaccharide is the dominant *O*-glycan at Ser$^{232}$. For Thr$^{236}$, the IgA1 myeloma protein and normal human serum IgA1 have similar distribution of *O*-glycan chains with the non-glycosylated form dominating followed by glycoforms with attached monosaccharide and disaccharide. Interestingly, the microheterogeneity distributions at Thr$^{233}$ differ. While an *O*-glycan chain is absent in >50% of all IgA1 glycoforms for both sources of IgA1, the IgA1 (Mce1) myeloma protein has a noticeably higher abundance of disaccharide at Thr$^{233}$.

## Discussion

This study reports the discovery of several isomeric *O*-glycoforms in an IgA1 myeloma protein and more importantly IgA1 isolated from normal human serum by use of high-resolution MS and AI-ECD tandem MS. In the process of localizing and defining the microheterogeneity of *O*-glycans at specific attachment sites, we also unambiguously confirmed a native site of *O*-glycosylation in the IgA1 HR at Thr$^{233}$.

Wada et al. recently reported the presence of an *O*-glycan chain at Thr$^{233}$ in IgA1 HR glycoform with GalNAc$_5$Gal$_4$ glycopeptides from healthy individuals [17]. Although they did not separate isomeric structures of IgA1 HR *O*-glycopeptides, a mixture of two isomeric structures based on alternative attachment of a GalNAc at Thr$^{233}$ or Thr$^{236}$ was detected by ETD tandem MS [17]. We have reported that this proteolytic fragment failed to produce sufficient ECD [10] and ETD [12] fragmentation for complete assignment of sites of *O*-glycan chains, most likely due to insufficient charge density for this longer proteolytic fragment [12, 48, 49]. In the present study, we confirm Thr$^{233}$ as the sixth site of native IgA1 HR *O*-glycosylation (Fig. 1E) in addition to the five common sites of *O*-glycosylation previously reported [5, 9, 10, 12]. Modification of Thr$^{233}$ was also detected in IgA1 from normal human serum (Fig. 3) in addition to the various structural isomers. Based on our label-free quantitative analysis, Gal-deficient GalNAc is attached more frequently at Thr$^{236}$ than at Thr$^{233}$ as would be expected based on the history of Thr$^{236}$ reported as a site of *O*-glycan attachment.

Glycosylation isomers are more traditionally thought of as being variations in glycosidic linkages; namely, *N*-glycan isomers attached to the same amino acid position. In the case of clustered *O*-glycans, the close proximity of adjacent Ser and Thr residues creates a pattern of *O*-glycan modifications that is dictated by multiple glycosyltransferases. These clustered modifications (or aberrant modifications) create epitopes for recognition by antibodies [28, 50] or other serum or cell-surface proteins. *In vitro* GalNAc-T activity experiments have demonstrated that addition of GalNAc to IgA1 HR synthetic peptides produces isomeric glycopeptides based on different sites of attachment. The discovery of these sets of naturally occurring amino acid positional isomers provides evidence that *in vivo O*-glycan synthesis produces the same result. Whether all IgA1 *O*-glycan chains are initiated by the same GalNAc-T enzyme remains to be seen. For the IgA1 myeloma protein, $Thr^{225}$, $Thr^{228}$ and $Ser^{232}$ were glycosylated predominantly by GalNAc-Gal disaccharide, whereas Gal-deficient GalNAc or the absence of glycan was determined at $Ser^{230}$, $Thr^{233}$ and $Thr^{236}$ in various IgA1 HR *O*-glycopeptides. The variability of *O*-glycan attachment at these three sites resulted in the observed structural isomers of IgA1 HR from IgA1 myeloma protein (Figs. 1 & 2). Our combination of on-line LC FT-ICR MS and AI-ECD FT-ICR tandem MS defines the absence of glycosylation, presence of monosaccharide or disaccharide, and thus the microheterogeneity of *O*-glycans at each individual site. Fig. 6 provides this compiled analysis for the IgA1 (Mce1) myeloma protein. The same amino acid positional isomers were also identified in two independent preparations of normal human serum IgA1. While the single-site microheterogeneity of the IgA1 isolated from sera of healthy controls closely mimics the model IgA1 (Mce1) myeloma protein results, a larger sample set should be analyzed before such a detailed summary can be made regarding normal serum IgA1 *O*-glycan microheterogeneity.

Gal-deficient IgA1 *O*-glycans have been recognized to play a role in the pathogenesis of IgA nephropathy (IgAN). Several studies by lectin ELISA, mass spectrometry and other methods have identified a greater abundance of Gal-deficient and asialo-IgA1 HR *O*-glycoforms from patients with IgAN vs. normal healthy controls. These aberrantly glycosylated *O*-glycans play a role in the formation of circulating immune complexes that are deposited in the glomeruli of the kidney. However, the specific nephritogenic IgA1 *O*-glycoforms in patients with IgAN have yet to be identified. It is highly possible that the nephritogenic IgA *O*-glycoforms arise not only from incomplete galactosylation and sialylation but also from attachment at differential sites due to changes in the expression of GalNAc-T enzymes.

Isomeric structures with NeuAc attached either in a linear fashion to Gal (GalNAc-Gal-NeuAc) or branched with attachment to GalNAc (GalNAc (NeuAc) -Gal) in monosialylated $Ser^{232}$-$Arg^{245}$ glycoforms were successfully identified by IRMPD FT-ICR tandem MS. Sialylated-T (GalNAc-Gal) and/or Tn (GalNAc) antigens are associated with several human diseases, such as cancer and some autoimmune diseases including IgAN [25, 50-55]. This combination of ECD-type fragmentation and IRMPD is useful for the analyses of such sialylated clustered *O*-glycopeptides. Glycan fragmentation analysis by use of IRMPD and collision induced association (CID) is an established method for delineation of isomers [56-58] including alternate sites of sialic acid attachment [59]. Parsing out isomers for individual sites within a clustered region of glycosylation makes this task even more difficult. Similar approaches that make use of LC tandem MS combinations (CID, IRMPD, and ECD/ETD) need to be further developed for the analysis of aberrant *N*-glycosylation of IgG in autoimmune and infectious disease, such as rheumatoid arthritis, lupus erythematosus, Crohn's disease, HIV-infection, and periodontal disease [60-67].

Throughout our previous studies, we have had the goal of completely defining the *O*-glycan heterogeneity of an IgA1 population from a single source. This includes the total range of *O*-glycoforms present [9, 12], the site-specific assignment of each *O*-glycan chain for each

glycoform [9, 12], and the relative distribution of each *O*-glycoforms in a single sample [12, 44]. If this could be attained, then the nephritogenic IgA1 *O*-glycoforms could be discerned within the context of a masking population of normally glycosylated IgA1. The advent of ECD fragmentation made this goal seem attainable although the specific combination of ECD/ETD-"friendly" proteolytic fragments was not immediately obvious [10, 12]. Our combination of overlapping IgA1 HR fragments also provide a means of creating a highly reproducible series of corroborating relative quantitative distributions [12]. To date, these *O*-glycan distributions are the most consistent as a desialylated series, likely due to the reduced complexity of the sample. We assume that each unique IgA1 HR proteolytic fragment has its own ionization efficiency. However, for a given IgA1 HR peptide, the addition of monosaccharide and disaccharide nonpolar *O*-glycan chains appears to have minimal affect on ionization efficiency so that relative distributions under the same LC/MS conditions are very consistent. In our current results it would appear that sialylation of clustered IgA1 *O*-glycans does not have as much suppressive affect as we thought (Supplemental Table 6). Also, it is important to note that our use of traditional RP C18 LC/MS creates an unbiased relative distribution profile that does not enrich for specific glycan structures as seen with lectin chromatography methods [68, 69]. Each distinct overlapping IgA1 HR glycopeptide is analyzed separately and the resulting relative distributions corroborate each other and create a total picture of IgA1 *O*-glycan heterogeneity at the monosaccharide and disaccharide level. In our previous analysis of a different IgA1 myeloma protein we demonstrated that clustered *O*-glycan localization by ECD and ETD fragmentation can be accomplished on the LC timescale with considerably less starting material than required for offline LC preparations. Our current results add a further requirement to insure that positional isomers are separated well enough to collect independent tandem MS spectra of each isomer.

Our results make the picture of IgA1 *O*-glycan heterogeneity better resolved by defining several combinations of naturally occurring amino acid positional isomers. Our approach using desialylated samples could be applied to other heavily *O*-glycosylated proteins such as mucins that have been implicated as biomarkers in a variety of cancers. Achieving complete heterogeneity analysis that includes amino acid positional information for sialylation is desirable and future developments in ECD/ETD fragmentation will most likely make this possible. Still, having completely defined *O*-glycan heterogeneity at the terminal GalNAc and galactosylated GalNAc levels provides the opportunity to address questions of glycosyltransferase(s) site-specificity. Changes in enzyme expression levels could produce aberrantly glycosylated proteins that serve as markers for altered cellular states. Our results indicate that the aberrancy could be due to changes in the position of *O*-glycan attachment resulting in structural isomers rather than a wholesale change in *O*-glycan distribution. The ability to delineate and characterize isomeric structures of *O*-glycoforms raises new questions about the role of amino acid positional isomeric variants. Does their distribution change under different cellular conditions or vary over time? Does the nature of the heterogeneity influence the severity of renal injury in IgAN? These and other questions can now be answered with this new level of analytical resolution.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

# References

1. Mestecky, J.; Moro, I.; Kerr, MA.; Woof, JM. Mucosal immunoglobulins. In: Mestecky, J.; Bienenstock, J.; Lamm, ME.; Mayer, L.; McGhee, JR.; Strober, W., editors. Mucosal Immunology. 3rd. Elsevier Academic Press; Amsterdam: 2005. p. 153-181.

2. Macpherson AJ, McCoy KD, Johansen FE, Brandtzaeg P. The immune geography of IgA induction and function. Mucosal Immunol. 2008; 1(1):11–22. [PubMed: 19079156]

3. Mestecky J, Russell MW. IgA subclasses. Monogr Allergy. 1986; 19:277–301. [PubMed: 3093850]

4. Baenziger J, Kornfeld S. Structure of the carbohydrate units of IgA1 immunoglobulin. II. Structure of the O-glycosidically linked oligosaccharide units. J Biol Chem. 1974; 249(22):7270–81. [PubMed: 4373463]

5. Mattu TS, Pleass RJ, Willis AC, Kilian M, Wormald MR, Lellouch AC, Rudd PM, Woof JM, Dwek RA. The glycosylation and structure of human serum IgA1, Fab, and Fc regions and the role of N-glycosylation on Fc alpha receptor interactions. J Biol Chem. 1998; 273(4):2260–72. [PubMed: 9442070]

6. Iwase H, Ishii-Karakasa I, Fujii E, Hotta K, Hiki Y, Kobayashi Y. Analysis of glycoform of O-glycan from human myeloma immunoglobulin A1 by gas-phase hydrazinolysis following pyridylamination of oligosaccharides. Anal Biochem. 1992; 206(1):202–5. [PubMed: 1280920]

7. Iwase H, Tanaka A, Hiki Y, Kokubo T, Ishii-Karakasa I, Kobayashi Y, Hotta K. Estimation of the number of O-linked oligosaccharides per heavy chain of human serum IgA1 by matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOFMS) analysis of the hinge glycopeptide. J Biochem. 1996; 120(2):393–7. [PubMed: 8889826]

8. Novak J, Tomana M, Kilian M, Coward L, Kulhavy R, Barnes S, Mestecky J. Heterogeneity of O-glycosylation in the hinge region of human IgA1. Mol Immunol. 2000; 37(17):1047–56. [PubMed: 11399322]

9. Renfrow MB, Cooper HJ, Tomana M, Kulhavy R, Hiki Y, Toma K, Emmett MR, Mestecky J, Marshall AG, Novak J. Determination of aberrant O-glycosylation in the IgA1 hinge region by electron capture dissociation fourier transform-ion cyclotron resonance mass spectrometry. J Biol Chem. 2005; 280(19):19136–45. [PubMed: 15728186]

10. Renfrow MB, Mackay CL, Chalmers MJ, Julian BA, Mestecky J, Kilian M, Poulsen K, Emmett MR, Marshall AG, Novak J. Analysis of O-glycan heterogeneity in IgA1 myeloma proteins by Fourier transform ion cyclotron resonance mass spectrometry: implications for IgA nephropathy. Anal Bioanal Chem. 2007; 389(5):1397–407. [PubMed: 17712550]

11. Wada Y, Dell A, Haslam SM, Tissot B, Canis K, Azadi P, Backstrom M, Costello CE, Hansson GC, Hiki Y, Ishihara M, Ito H, Kakehi K, Karlsson N, Hayes CE, Kato K, Kawasaki N, Khoo KH, Kobayashi K, Kolarich D, Kondo A, Lebrilla C, Nakano M, Narimatsu H, Novak J, Novotny MV, Ohno E, Packer NH, Palaima E, Renfrow MB, Tajiri M, Thomsson KA, Yagi H, Yu SY, Taniguchi N. Comparison of methods for profiling O-glycosylation: Human Proteome Organisation Human Disease Glycomics/Proteome Initiative multi-institutional study of IgA1. Mol Cell Proteomics. 2010; 9(4):719–27. [PubMed: 20038609]

12. Takahashi K, Wall SB, Suzuki H, Smith AD, Hall S, Paulsen K, Kilian M, Mobley JA, Julian BA, Mestecky J, Novak J, Renfrow MB. Clustered O-glycans of IgA1: Defining macro- and micro-heterogeneity by use of electron capture/transfer dissociation. Mol Cell Proteomics. 2010; 9(11): 2545–2557. [PubMed: 20823119]

13. Iwase H, Tanaka A, Hiki Y, Kokubo T, Ishii-Karakasa I, Nishikido J, Kobayashi Y, Hotta K. Application of matrix-assisted laser desorption ionization time-of-flight mass spectrometry to the analysis of glycopeptide-containing multiple O-linked oligosaccharides. J Chromatogr B Biomed Sci Appl. 1998; 709(1):145–9. [PubMed: 9653936]

14. Hiki Y, Tanaka A, Kokubo T, Iwase H, Nishikido J, Hotta K, Kobayashi Y. Analyses of IgA1 hinge glycopeptides in IgA nephropathy by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. J Am Soc Nephrol. 1998; 9(4):577–82. [PubMed: 9555659]

15. Iwase H, Tanaka A, Hiki Y, Kokubo T, Sano T, Ishii-Karakasa I, Hisatani K, Kobayashi Y, Hotta K. Analysis of the microheterogeneity of the IgA1 hinge glycopeptide having multiple O-linked oligosaccharides by capillary electrophoresis. Anal Biochem. 2001; 288(1):22–7. [PubMed: 11141302]

16. Tarelli E, Smith AC, Hendry BM, Challacombe SJ, Pouria S. Human serum IgA1 is substituted with up to six O-glycans as shown by matrix assisted laser desorption ionisation time-of-flight mass spectrometry. Carbohydr Res. 2004; 339(13):2329–35. [PubMed: 15337464]

17. Wada Y, Tajiri M, Ohshima S. Quantitation of saccharide compositions of O-glycans by mass spectrometry of glycopeptides and its application to rheumatoid arthritis. J Proteome Res. 2010; 9(3):1367–73. [PubMed: 20104905]

18. Odani H, Hiki Y, Takahashi M, Nishimoto A, Yasuda Y, Iwase H, Shinzato T, Maeda K. Direct evidence for decreased sialylation and galactosylation of human serum IgA1 Fc O-glycosylated hinge peptides in IgA nephropathy by mass spectrometry. Biochem Biophys Res Commun. 2000; 271(1):268–74. [PubMed: 10777713]

19. Hiki Y, Odani H, Takahashi M, Yasuda Y, Nishimoto A, Iwase H, Shinzato T, Kobayashi Y, Maeda K. Mass spectrometry proves under-O-glycosylation of glomerular IgA1 in IgA nephropathy. Kidney Int. 2001; 59(3):1077–85. [PubMed: 11231363]

20. Shimozato S, Hiki Y, Odani H, Takahashi K, Yamamoto K, Sugiyama S. Serum under-galactosylated IgA1 is increased in Japanese patients with IgA nephropathy. Nephrol Dial Transplant. 2008; 23(6):1931–9. [PubMed: 18178603]

21. Iwasaki H, Zhang Y, Tachibana K, Gotoh M, Kikuchi N, Kwon YD, Togayachi A, Kudo T, Kubota T, Narimatsu H. Initiation of O-glycan synthesis in IgA1 hinge region is determined by a single enzyme, UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 2. J Biol Chem. 2003; 278(8):5613–21. [PubMed: 12438318]

22. Wandall HH, Irazoqui F, Tarp MA, Bennett EP, Mandel U, Takeuchi H, Kato K, Irimura T, Suryanarayanan G, Hollingsworth MA, Clausen H. The lectin domains of polypeptide GalNAc-transferases exhibit carbohydrate-binding specificity for GalNAc: lectin binding to GalNAc-glycopeptide substrates is required for high density GalNAc-O-glycosylation. Glycobiology. 2007; 17(4):374–87. [PubMed: 17215257]

23. Tarp MA, Clausen H. Mucin-type O-glycosylation and its potential use in drug and vaccine development. Biochim Biophys Acta. 2008; 1780(3):546–63. [PubMed: 17988798]

24. Suzuki H, Kiryluk K, Novak J, Moldoveanu Z, Herr AB, Renfrow MB, Wyatt RJ, Scolari F, Mestecky J, Gharavi AG, Julian BA. The pathophysiology of IgA nephropathy. J Am Soc Nephrol. 2011; 22(10):1795–803. [PubMed: 21949093]

25. Suzuki H, Moldoveanu Z, Hall S, Brown R, Vu HL, Novak L, Julian BA, Tomana M, Wyatt RJ, Edberg JC, Alarcon GS, Kimberly RP, Tomino Y, Mestecky J, Novak J. IgA1-secreting cell lines from patients with IgA nephropathy produce aberrantly glycosylated IgA1. J Clin Invest. 2008; 118(2):629–39. [PubMed: 18172551]

26. Novak J, Julian BA, Tomana M, Mestecky J. IgA glycosylation and IgA immune complexes in the pathogenesis of IgA nephropathy. Semin Nephrol. 2008; 28(1):78–87. [PubMed: 18222349]

27. Tomana M, Matousovic K, Julian BA, Radl J, Konecny K, Mestecky J. Galactose-deficient IgA1 in sera of IgA nephropathy patients is present in complexes with IgG. Kidney Int. 1997; 52(2):509–16. [PubMed: 9264010]

28. Tomana M, Novak J, Julian BA, Matousovic K, Konecny K, Mestecky J. Circulating immune complexes in IgA nephropathy consist of IgA1 with galactose-deficient hinge region and antiglycan antibodies. J Clin Invest. 1999; 104(1):73–81. [PubMed: 10393701]

29. Moldoveanu Z, Wyatt RJ, Lee JY, Tomana M, Julian BA, Mestecky J, Huang WQ, Anreddy SR, Hall S, Hastings MC, Lau KK, Cook WJ, Novak J. Patients with IgA nephropathy have increased serum galactose-deficient IgA1 levels. Kidney Int. 2007; 71(11):1148–54. [PubMed: 17342176]

30. Allen AC, Bailey EM, Brenchley PE, Buck KS, Barratt J, Feehally J. Mesangial IgA1 in IgA nephropathy exhibits aberrant O-glycosylation: observations in three patients. Kidney Int. 2001; 60(3):969–73. [PubMed: 11532091]

31. Andre PM, Le Pogamp P, Chevet D. Impairment of jacalin binding to serum IgA in IgA nephropathy. J Clin Lab Anal. 1990; 4(2):115–9. [PubMed: 2313468]

32. Mestecky J, Tomana M, Crowley-Nowick PA, Moldoveanu Z, Julian BA, Jackson S. Defective galactosylation and clearance of IgA1 molecules as a possible etiopathogenic factor in IgA nephropathy. Contrib Nephrol. 1993; 104:172–82. [PubMed: 8325028]

33. Allen AC. Abnormal glycosylation of IgA: is it related to the pathogenesis of IgA nephropathy? Nephrol Dial Transplant. 1995; 10(7):1121–4. [PubMed: 7478108]

34. Smith AC, Molyneux K, Feehally J, Barratt J. O-glycosylation of serum IgA1 antibodies against mucosal and systemic antigens in IgA nephropathy. J Am Soc Nephrol. 2006; 17(12):3520–8. [PubMed: 17093066]

35. Lindfors K, Suzuki H, Novak J, Collin P, Saavalainen P, Koskinen LL, Maki M, Kaukinen K. Galactosylation of serum IgA1 O-glycans in celiac disease. J Clin Immunol. 2011; 31(1):74–9. [PubMed: 20938724]

36. Dechant M, Beyer T, Schneider-Merck T, Weisner W, Peipp M, van de Winkel JG, Valerius T. Effector mechanisms of recombinant IgA antibodies against epidermal growth factor receptor. J Immunol. 2007; 179(5):2936–43. [PubMed: 17709508]

37. Dechant M, Vidarsson G, Stockmeyer B, Repp R, Glennie MJ, Gramatzki M, van De Winkel JG, Valerius T. Chimeric IgA antibodies against HLA class II effectively trigger lymphoma cell killing. Blood. 2002; 100(13):4574–80. [PubMed: 12393717]

38. Ashline DJ, Lapadula AJ, Liu YH, Lin M, Grace M, Pramanik B, Reinhold VN. Carbohydrate structural isomers analyzed by sequential mass spectrometry. Anal Chem. 2007; 79(10):3830–42. [PubMed: 17397137]

39. Aoki K, Porterfield M, Lee SS, Dong B, Nguyen K, McGlamry KH, Tiemeyer M. The diversity of O-linked glycans expressed during Drosophila melanogaster development reflects stage-and tissue-specific requirements for cell signaling. J Biol Chem. 2008; 283(44):30385–400. [PubMed: 18725413]

40. Prien JM, Huysentruyt LC, Ashline DJ, Lapadula AJ, Seyfried TN, Reinhold VN. Differentiating N-linked glycan structural isomers in metastatic and nonmetastatic tumor cells using sequential mass spectrometry. Glycobiology. 2008; 18(5):353–66. [PubMed: 18256178]

41. Mestecky J, Kilian M. Immunoglobulin A (IgA). Methods Enzymol. 1985; 116:37–75. [PubMed: 3911017]

42. Moore JS, Kulhavy R, Tomana M, Moldoveanu Z, Suzuki H, Brown R, Hall S, Kilian M, Poulsen K, Mestecky J, Julian BA, Novak J. Reactivities of N-acetylgalactosamine-specific lectins with human IgA1 proteins. Mol Immunol. 2007; 44(10):2598–604. [PubMed: 17275907]

43. Sullivan B, Addona TA, Carr SA. Selective detection of glycopeptides on ion trap massspectrometers. Anal Chem. 2004; 76(11):3112–8. [PubMed: 15167790]

44. Rebecchi KR, Wenke JL, Go EP, Desaire H. Label-free quantitation: a new glycoproteomics approach. J Am Soc Mass Spectrom. 2009; 20(6):1048–59. [PubMed: 19278867]

45. Perdivara I, Petrovich R, Allinquant B, Deterding LJ, Tomer KB, Przybylski M. Elucidation of O-glycosylation structures of the beta-amyloid precursor protein by liquid chromatography-mass spectrometry using electron transfer dissociation and collision induced dissociation. J Proteome Res. 2009; 8(2):631–42. [PubMed: 19093876]

46. Hakansson K, Cooper HJ, Emmett MR, Costello CE, Marshall AG, Nilsson CL. Electron capture dissociation and infrared multiphoton dissociation MS/MS of an N-glycosylated tryptic peptic to yield complementary sequence information. Anal Chem. 2001; 73(18):4530–6. [PubMed: 11575803]

47. Hakansson K, Chalmers MJ, Quinn JP, McFarland MA, Hendrickson CL, Marshall AG. Combined electron capture and infrared multiphoton dissociation for multistage MS/MS in a Fourier transform ion cyclotron resonance mass spectrometer. Anal Chem. 2003; 75(13):3256–62. [PubMed: 12964777]

48. Darula Z, Medzihradszky KF. Affinity enrichment and characterization of mucin core-1 type glycopeptides from bovine serum. Mol Cell Proteomics. 2009; 8(11):2515–26. [PubMed: 19674964]

49. Good DM, Wirtala M, McAlister GC, Coon JJ. Performance characteristics of electron transfer dissociation mass spectrometry. Mol Cell Proteomics. 2007; 6(11):1942–51. [PubMed: 17673454]

50. Suzuki H, Fan R, Zhang Z, Brown R, Hall S, Julian BA, Chatham WW, Suzuki Y, Wyatt RJ, Moldoveanu Z, Lee JY, Robinson J, Tomana M, Tomino Y, Mestecky J, Novak J. Aberrantly glycosylated IgA1 in IgA nephropathy patients is recognized by IgG antibodies with restricted heterogeneity. J Clin Invest. 2009; 119(6):1668–77. [PubMed: 19478457]

51. Sames D, Chen XT, Danishefsky SJ. Convergent total synthesis of a tumour-associated mucin motif. Nature. 1997; 389(6651):587–91. [PubMed: 9335496]

52. Dall'Olio F, Chiricolo M. Sialyltransferases in cancer. Glycoconj J. 2001; 18(11–12):841–50. [PubMed: 12820717]

53. Marcos NT, Pinho S, Grandela C, Cruz A, Samyn-Petit B, Harduin-Lepers A, Almeida R, Silva F, Morais V, Costa J, Kihlberg J, Clausen H, Reis CA. Role of the human ST6GalNAc-I and ST6GalNAc-II in the synthesis of the cancer-associated sialyl-Tn antigen. Cancer Res. 2004; 64(19):7050–7. [PubMed: 15466199]

54. van Schooten CJ, Denis CV, Lisman T, Eikenboom JC, Leebeek FW, Goudemand J, Fressinaud E, van den Berg HM, de Groot PG, Lenting PJ. Variations in glycosylation of von Willebrand factor with O-linked sialylated T antigen are associated with its plasma levels. Blood. 2007; 109(6): 2430–7. [PubMed: 17090649]

55. Ohyabu N, Hinou H, Matsushita T, Izumi R, Shimizu H, Kawamoto K, Numata Y, Togame H, Takemoto H, Kondo H, Nishimura S. An essential epitope of anti-MUC1 monoclonal antibody KL-6 revealed by focused glycopeptide library. J Am Chem Soc. 2009; 131(47):17102–9. [PubMed: 19899793]

56. Medzihradszky KF, Gillece-Castro BL, Settineri CA, Townsend RR, Masiarz FR, Burlingame AL. Structure determination of O-linked glycopeptides by tandem mass spectrometry. Biomed Environ Mass Spectrom. 1990; 19(12):777–81. [PubMed: 1708302]

57. Medzihradszky KF, Gillece-Castro BL, Townsend RR, Burlingame AL, Hardy MR. Structural elucidation of O-linked glycopeptides by high energy collision-induced dissociation. J Amer Soc Mass Spect. 1996; 7(4):319–328.

58. Zaia J. Mass spectrometry of oligosaccharides. Mass Spectrom Rev. 2004; 23(3):161–227. [PubMed: 14966796]

59. Sagi D, Peter-Katalinic J, Conradt HS, Nimtz M. Sequencing of tri- and tetraantennary N-glycans containing sialic acid by negative mode ESI QTOF tandem MS. J Am Soc Mass Spectrom. 2002; 13(9):1138–48. [PubMed: 12322961]

60. Arnold JN, Wormald MR, Sim RB, Rudd PM, Dwek RA. The impact of glycosylation on the biological function and structure of human immunoglobulins. Annu Rev Immunol. 2007; 25:21–50. [PubMed: 17029568]

61. Malhotra R, Wormald MR, Rudd PM, Fischer PB, Dwek RA, Sim RB. Glycosylation changes of IgG associated with rheumatoid arthritis can activate complement via the mannose-binding protein. Nat Med. 1995; 1(3):237–43. [PubMed: 7585040]

62. Moore JS, Wu X, Kulhavy R, Tomana M, Novak J, Moldoveanu Z, Brown R, Goepfert PA, Mestecky J. Increased levels of galactose-deficient IgG in sera of HIV-1-infected individuals. AIDS. 2005; 19(4):381–9. [PubMed: 15750391]

63. Novak J, Tomana M, Shah GR, Brown R, Mestecky J. Heterogeneity of IgG glycosylation in adult periodontal disease. J Dent Res. 2005; 84(10):897–901. [PubMed: 16183787]

64. Rademacher TW, Parekh RB, Dwek RA, Isenberg D, Rook G, Axford JS, Roitt I. The role of IgG glycoforms in the pathogenesis of rheumatoid arthritis. Springer Semin Immunopathol. 1988; 10(2–3):231–49. [PubMed: 3055379]

65. Tomana M, Schrohenloher RE, Reveille JD, Arnett FC, Koopman WJ. Abnormal galactosylation of serum IgG in patients with systemic lupus erythematosus and members of families with high frequency of autoimmune diseases. Rheumatol Int. 1992; 12(5):191–4. [PubMed: 1290021]

66. Troelsen LN, Garred P, Madsen HO, Jacobsen S. Genetically determined high serum levels of mannose-binding lectin and agalactosyl IgG are associated with ischemic heart disease in rheumatoid arthritis. Arthritis Rheum. 2007; 56(1):21–9. [PubMed: 17195187]

67. Youings A, Chang SC, Dwek RA, Scragg IG. Site-specific glycosylation of human immunoglobulin G is altered in four rheumatoid arthritis patients. Biochem J. 1996; 314(Pt 2): 621–30. [PubMed: 8670078]

68. Durham M, Regnier FE. Targeted glycoproteomics: serial lectin affinity chromatography in the selection of O-glycosylation sites on proteins from the human blood proteome. J Chromatogr A. 2006; 1132(1–2):165–73. [PubMed: 16919642]

69. Schwientek T, Mandel U, Roth U, Muller S, Hanisch FG. A serial lectin approach to the mucin-type O-glycoproteome of Drosophila melanogaster S2 cells. Proteomics. 2007; 7(18):3264–77. [PubMed: 17708590]
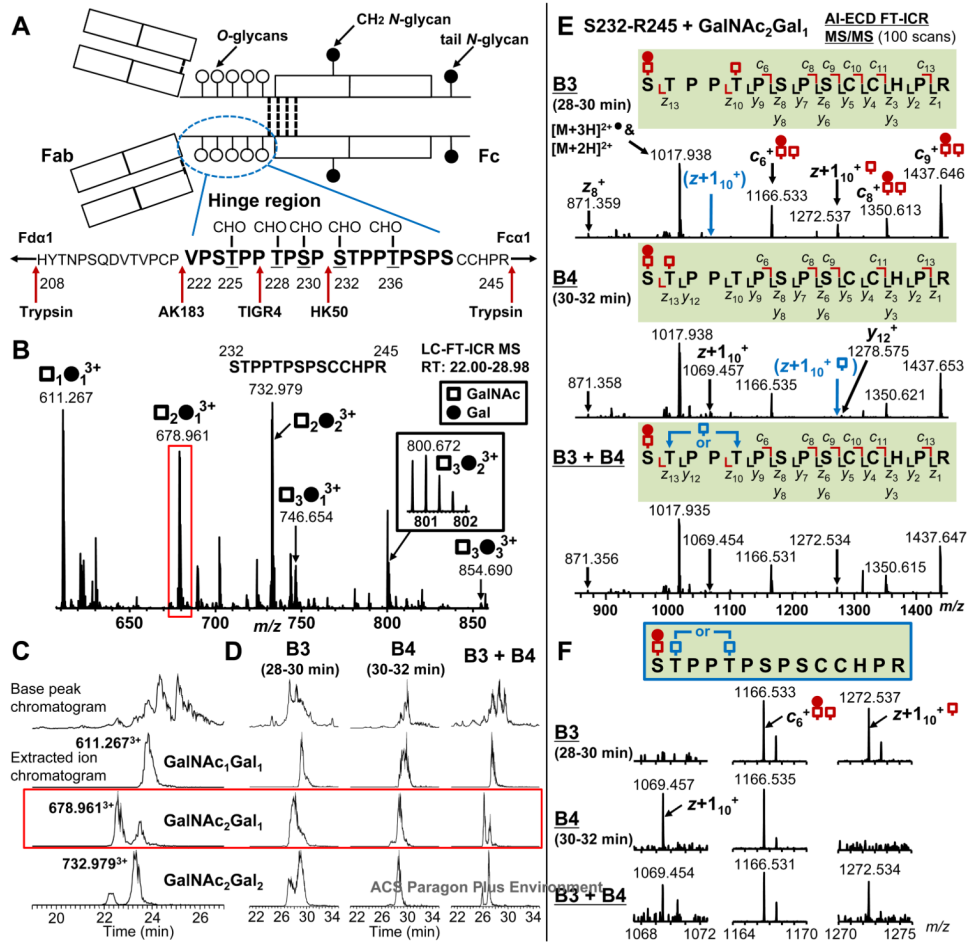
**Fig. 1.**

Identification of structural isomers in IgA1 hinge-region (HR) *O*-glycopeptides. (*A*). IgA1 HR structure. The HR is a Pr*O*-rich segment with nine possible sites of *O*-glycan attachment. Underlined Ser and Thr residues are glycosylated including Thr[233] [5] Red arrows show cleavage sites of trypsin and three IgA-specific proteases (from *Clostridium ramosum* strain AK183, *Streptococcus pneumoniae* strain TIGR4 or *Haemophilus influenzae* strain HK50). (*B*). Online LC FT-ICR MS analysis of Ser[232]-Arg[245] HR glycopeptides of IgA1 (Mce1) myeloma protein. The number of *O*-glycan chains was assigned based on the masses of the amino-acid sequence, GalNAc (open squares), and Gal (full circles). The *m/z* values for the observed glycopeptides, relative abundance, and mass error are summarized in Table 1. (*C*). LC-XIC of three dominant glycopeptides in the C-terminal HR fragments (Ser[232]-Arg[245]) were individually extracted for the specific *m/z* of glycopeptides. A single peak was observed for the GalNAc$_1$Gal$_1$ Ser[232]-Arg[245] glycoform, whereas bimodal peaks were observed for GalNAc$_2$Gal$_1$ and GalNAc$_2$Gal$_2$ Ser[232]-Arg[245] glycopeptides. (*D*). XIC of Ser[232]-Arg[245] glycopeptides with GalNAc$_1$Gal$_1$, GalNAc$_2$Gal$_1$, and GalNAc$_2$Gal$_2$ derived from IgA1 (Mce1) myeloma protein and fractionated by offline LC. The offline LC fractions from 28-30 min (B3) and 30-32 (B4), and their mixtures (B3+B4) (1:1 v/v) were re-analyzed by online LC FT-ICR MS and the XIC of the corresponding glycopeptide ions were obtained. (*E*). AI-ECD FT-ICR tandem MS of the IgA1 HR Ser[232]-Arg[245] + GalNAc$_2$Gal$_1$ isomers (red rectangle shown in C and D) in separate fractions. AI-ECD fragmentation was performed using a triply charged precursor ion. The absence of the key *z*+1$_{10}$ fragment which indicated isomeric structure is highlighted in blue. The observed *c*, *z*, and y fragments

for each *O*-glycopeptide are indicated above and below the sequences. (*F*). Distinguishing fragment ions of structural isomers of the $GalNAc_2Gal_1$ $Ser^{232}$-$Arg^{245}$ glycoform. Arrows in the spectra denote the key fragments that allowed assignment of glycan attachment sites. The alternative attachment sites in the isomeric structural variants of glycopeptides were shown in the amino-acid sequence.
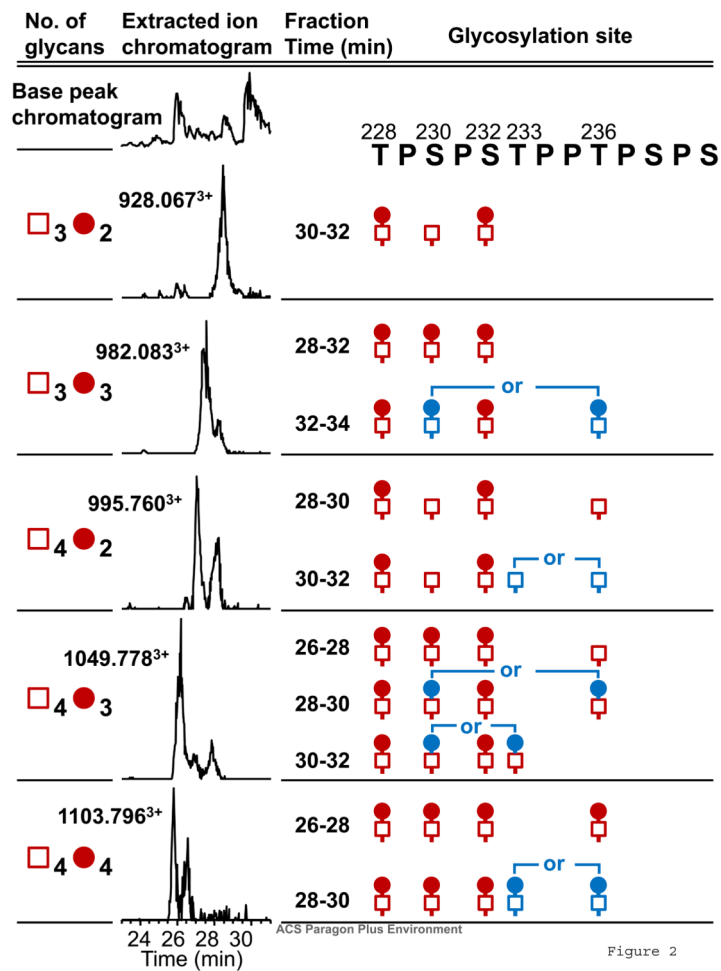
**Fig. 2.**
Summary of Thr$^{228}$-Arg$^{245}$ glycoforms of IgA1 (Mce1) myeloma protein. The glycopeptides were separated into three microfractionated by offline LC and individually analyzed by AI-ECD tandem MS to determine the sites of *O*-glycan attachment. Glycans in blue color indicate the alternative attachment sites in the isomeric variants of glycopeptides. All fragmentation spectra are shown in supplemental data.
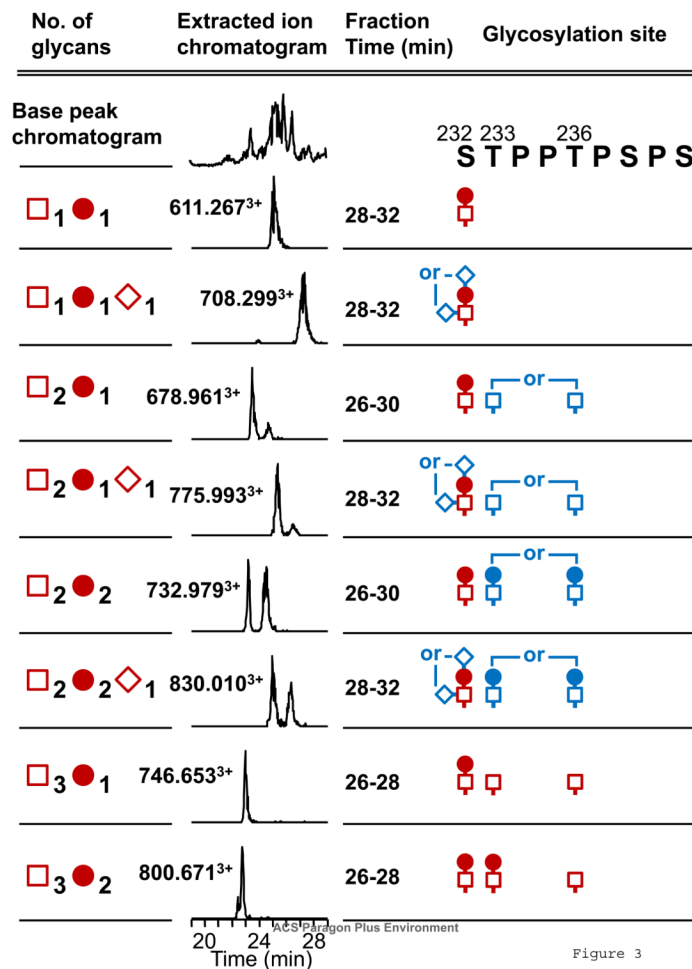
**Fig. 3.**
Summary of Ser$^{232}$-Arg$^{245}$ glycoforms of normal human serum IgA1. The microfractionated glycopeptides were individually analyzed by AI-ECD tandem MS to identify the sites of glycan attachment. Glycans in blue color indicate the alternative attachment sites in the isomeric structural variants of glycopeptides. All fragmentation spectra are shown in supplemental data.
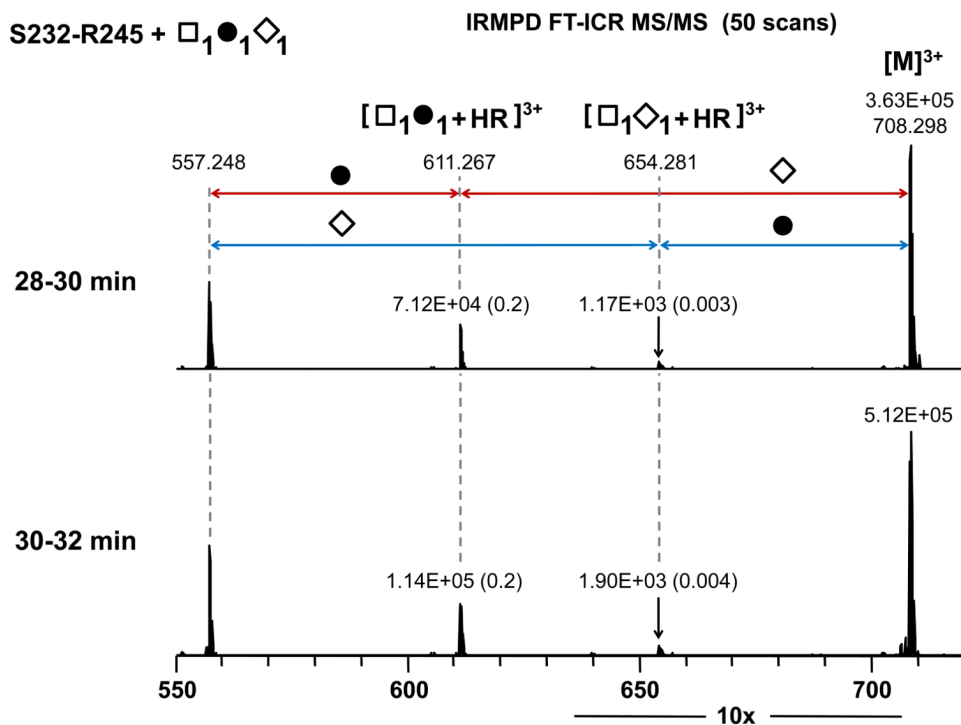
**Fig. 4.**
IRMPD FT-ICR tandem MS of the normal human serum IgA1 HR $Ser^{232}$-$Arg^{245}$ with $GalNAc_1Gal_1NeuAc_1$. Red fragmentation pathway indicates NeuAc (open diamond) attachment to Gal (filled circle), whereas blue fragmentation pathway shows GalNAc (open rectangle) with NeuAc.
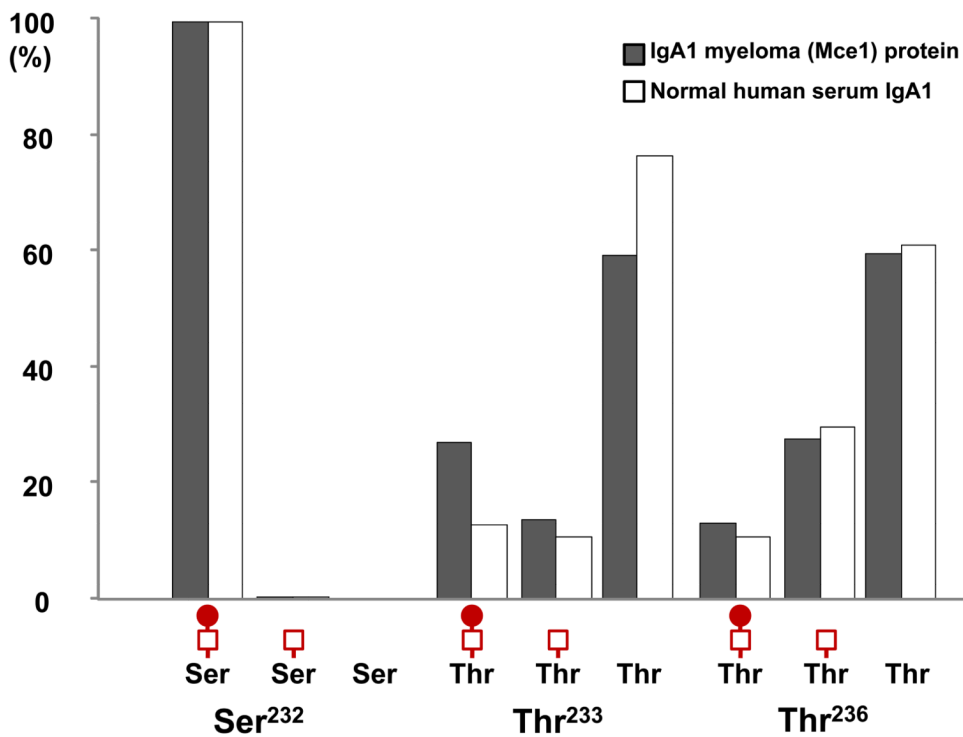
**Fig. 5.**
Comparison of individual site microheterogeneity for Ser[232], Thr[233], and Thr[236] between the model IgA1 (Mce1) myeloma protein and IgA1 isolated from sera of healthy controls. Ser[232] is dominated by a GalNAc-Gal disaccharide. Thr[233] and Thr[236] in both samples show more single-site heterogeneity with a noticeable difference in the amount of disaccharide present at Thr[233].

| | Relative abundance: microheterogeneity at individual sites (%)[a] | | | | | |
|---|---|---|---|---|---|---|
| | Thr[225] | Thr[228] | Ser[230] | Ser[232] | Thr[233] | Thr[236] |
| Not glycosylated | 0.11 ± 0.05 | N.D. | 11.96 ± 0.56 | N.D. | 59.41 ±1.18 | 59.49 ±1.50 |
| Gal-deficient | 0.24 ± 0.02 | 1.11 ± 0.08 | 31.78 ± 0.85 | 0.48 ± 0.06 | 13.59 ±0.53 | 27.53 ±1.01 |
| Disaccharide | 99.65 ± 0.05 | 98.89 ± 0.08 | 56.26 ± 0.45 | 99.52 ± 0.06 | 27.00 ±0.72 | 12.98 ±0.53 |

**Fig. 6.**

The *O*-glycan microheterogeneity of each site for IgA1 (Mce1) myeloma protein. Ser[230], Thr[233] and Thr[236] are the dominant sites of *O*-glycan variability, including non-glycosylated as well as Gal-deficient O-glycans. Thr[225], Thr[228] and Ser[232] were predominantly modified by a disaccharide.[a] Relative abundance (mean ± S.D. in triplicate of LC/MS runs) of individual site microheterogeneity is expressed as percentage against the sum of total XIC of all glycopeptides.

**Table 1**

**Identified HR *O*-glycopeptides of IgA1 (Mce1) myeloma protein**

| Glycan structure | $m/z$[a] | Retention time (min) | Relative abundance (%)[b] | Mass error (ppm) | AI-ECD MS/MS | Site-specific *O*-glycosylation | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | Ser[232] | Thr[233] | Thr[236] |
| **Peptide sequence:[232] STPPTPSPSCCHPR[245]** | | | | | | | | |
| GalNAc$_1$ | 557.249 | 25.83 - 27.14 | 0.25 ± 0.05 | 0.57 | | | | |
| GalNAc$_1$Gal$_1$ | 611.267 | 25.21 - 28.78 | 30.32 ± 0.66 | 1.89 | d | • | | |
| GalNAc$_2$ | 624.943 | 26.27 - 26.82 | 0.13 ± 0.04 | 2.70 | | | | |
| GalNAc$_2$Gal$_1$ | 678.961 | 23.66 - 26.69 | 24.43 ± 1.63 | 2.98 | d | | | |
| *1st peak*[c] | | 23.66 - 24.85 | *16.28 ± 1.39* | | d | • | | |
| *2nd peak*[c] | | 24.85 - 26.69 | *8.15 ± 0.82* | | d | • | | |
| GalNAc$_2$Gal$_2$ | 732.979 | 23.28 - 26.19 | 33.22 ± 0.82 | 2.71 | d | | | |
| *1st peak*[c] | | 23.28 - 24.52 | *12.44 ± 0.36* | | d | | | • |
| *2nd peak*[c] | | 24.52 - 26.19 | *20.78 ± 1.09* | | d | | • | |
| GalNAc$_3$ | 692.638 | 24.41 - 24.79 | 0.10 ± 0.06 | 4.46 | | • | | |
| GalNAc$_3$Gal$_1$ | 746.654 | 23.28 - 24.56 | 5.34 ± 0.87 | 2.57 | d | • | | |
| GalNAc$_3$Gal$_2$ | 800.672 | 22.82 - 24.45 | 5.68 ± 0.69 | 3.19 | d | • | • | |
| GalNAc$_3$Gal$_3$ | 854.690 | 22.51 - 23.32 | 0.54 ± 0.25 | 3.38 | | | | |

[a] Glycopeptides were ionized as 3+ charged ions.

[b] Relative abundance of each glycopeptide is expressed as percentage against the sum of XIC of all glycopeptides with the same amino-acid sequence, as described 12, 44. Complete list of glycopeptides is available in Supplemental Table 1.

[c] Two XIC peaks were identified, as shown in Fig. 1C.

[d] All AI-ECD MS/MS spectra are available in Supplemental data. GalNAc: open squares, Gal: filled circles.

**Table 2**

**Identified HR isomeric *O*-glycopeptides of IgA1 from normal human serum**

| Glycan structure | $m/z$[a] | Retention time (min) | Relative abundance (%)[b] | Mass error (ppm) | AI-ECD MS/MS | Site-specific *O*-glycosylation | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | Ser [232] | Thr [233] | Thr [236] |
| Peptide sequence: [232]STPPTPSPSCCHPR [245] | | | | | | | | |
| GalNAc₁Gal₁ | 611.267 | 24.77 - 26.38 | 3.04 ± 0.62 | 1.14 | ✓[d] | -• | | |
| GalNAc₁Gal₁NeuAc₁ | 708.299 | 25.18 - 30.20 | 36.82 ± 1.67 | 1.55 | ✓[d] | -• | | |
| GalNAc₂Gal₁ | 678.961 | 23.25 - 25.94 | 3.22 ± 0.94 | 2.45 | ✓[d] | | | |
| *1st peak*[c] | | *23.25-24.31* | *2.27 ± 0.79* | | ✓[d] | -• | | |
| *2nd peak*[c] | | *24.31-25.94* | *0.95 ± 0.16* | | ✓[d] | -• | | |
| GalNAc₂Gal₁NeuAc₁ | 775.993 | 24.58 - 27.94 | 23.49 ± 1.36 | 2.19 | ✓[d] | | | |
| *1st peak*[c] | | *24.58-25.91* | *18.52 ± 0.73* | | ✓[d] | -• | | |
| *2nd peak*[c] | | *25.91-27.94* | *4.97 ± 0.66* | | ✓[d] | -• | | |
| GalNAc₂Gal₂ | 732.979 | 22.96 - 25.35 | 2.91 ± 0.43 | 3.05 | ✓[d] | | | |
| *1st peak*[c] | | *22.96-23.90* | *1.09 ± 0.26* | | ✓[d] | -• | | -• |
| *2nd peak*[c] | | *23.90-25.35* | *1.82 ± 0.20* | | ✓[d] | -• | -• | |
| GalNAc₂Gal₂NeuAc₁ | 830.010 | 24.13 - 28.15 | 11.00 ± 0.97 | 1.72 | ✓[d] | | | |
| *1st peak*[c] | | *24.13-25.53* | *5.64 ± 0.24* | | ✓[d] | -• | | -• |
| *2nd peak*[c] | | *25.53-28.15* | *5.37 ± 1.03* | | ✓[d] | -• | -• | |
| GalNAc₃Gal₁ | 746.653 | 22.74 - 24.07 | 1.74 ± 0.24 | 1.43 | ✓[d] | -• | | |
| GalNAc₃Gal₂ | 800.671 | 22.15 - 23.49 | 0.65 ± 0.20 | 1.12 | ✓[d] | -• | -• | |

[a] Glycopeptides are ionized as 3+ charged ions.

[b] Relative abundance of each glycopeptide is expressed as percentage against the sum of XIC of all glycopeptides with the same amin-acid sequence, as described [12, 44]. Whole list of identified glycopeptides is available in Supplemental table 4.

[c] Two XIC peaks were identified, as shown in Fig. 4.

[d] All AI-ECD MS/MS spectra are available in Supplemental data.

GalNAc: open squares, Gal: filled circles, NeuAc: open diamond.