



Universiteit  
Leiden  
The Netherlands

## **Unravelling the sugar-coating of prostate-specific antigen : method development and its application to prostate cancer research**

Kammeijer, G.S.M.

### **Citation**

Kammeijer, G. S. M. (2019, March 7). *Unravelling the sugar-coating of prostate-specific antigen : method development and its application to prostate cancer research*. Retrieved from <https://hdl.handle.net/1887/69482>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/69482>

**Note:** To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/69482> holds various files of this Leiden University dissertation.

**Author:** Kammeijer, G.S.M.

**Title:** Unravelling the sugar-coating of prostate-specific antigen : method development and its application to prostate cancer research

**Issue Date:** 2019-03-07



# **SIALIC ACID LINKAGE DIFFERENTIATION OF GLYCOPEPTIDES**

**USING CAPILLARY ELECTROPHORESIS -  
ELECTROSPRAY IONIZATION -  
MASS SPECTROMETRY**

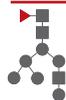
Based on Guinevere S.M. Kammeijer, Bas C. Jansen, Isabelle Kohler, Anthonius A.M. Heemskerk, Oleg A. Mayboroda, Paul J. Hensbergen, Julie Schappler, Manfred Wuhrer, "*Sialic Acid Linkage Differentiation of Glycopeptides using Capillary Electrophoresis - Electrospray Ionization - Mass Spectrometry*", Scientific Reports 7 (1), Article number: 3733, 2017.



## ABSTRACT

Sialylation is a glycosylation feature that occurs in different linkages at the non-reducing end of a glycan moiety, the linkage isomers are often differentially associated with various biological processes. Due to very similar physico-chemical properties, the separation of isomeric sialylated glycopeptides remains challenging but of utmost importance in the biomedicine and biotechnology, including biomarker discovery, glyco-engineering and biopharmaceutical characterization. This study presents the implementation of a high-resolution separation platform based on capillary electrophoresis - electrospray ionization - mass spectrometry (CE-ESI-MS) allowing for the selective analysis of  $\alpha$ 2,3- and  $\alpha$ 2,6-sialylated glycopeptides. These differentially linked glycopeptides showed an identical fragmentation pattern (collision induced dissociation) but different electrophoretic mobilities, allowing for baseline separation of the different linkages without the need for an extensive sample preparation. The different migration behavior between the two moieties was found to correlate with differences in  $pK_a$  values. Using a novel methodology adapted from the so-called internal standard CE approach, a relative difference of  $3.4 \cdot 10^{-2}$  in  $pK_a$  unit was determined. This approach was applied for the analysis of tryptic glycopeptides of prostate-specific antigen, which shows highly complex and heterogeneous glycosylation. The developed platform therefore appears attractive for the identification of differentially linked sialic acids that may be related to pathological conditions.

2





## INTRODUCTION

Glycosylation is one of the most complex post-translational modifications of proteins that affects protein folding, stability, half-life, protein-protein interactions, signaling and trafficking.<sup>133-135</sup> Carbohydrate moieties are attached to the peptide backbone either via the nitrogen atom of the side-chain of asparagine (*N*-linked) with the sequence *N* - *X* - *S/T* (*X* ≠ *P*) or via an oxygen atom (*O*-linked) of serine or threonine. Next to the diversity in site occupancy (macroheterogeneity), a large variety in glycan structures may also be observed at a given glycosylation site (microheterogeneity).

Changes in glycosylation may be associated with various factors, including age and pathophysiological conditions.<sup>136-139</sup> In certain disease settings, specific glycan features such as fucosylation, bisection, galactosylation or sialylation are known to be affected.<sup>140,141</sup> Sialylation, *i.e.*, a process by which sialic acids are attached to glycans, occurs on different monosaccharides and in different linkage positions, namely  $\alpha$ 2,3,  $\alpha$ 2,6 or  $\alpha$ 2,8. The motifs are located at the non-reducing end of the glycans, creating binding sites for multiple human lectins, toxins and pathogens.<sup>82</sup> Sialic acids are involved in numerous pathophysiological processes. For instance,  $\alpha$ 2,3-linked sialic acids have shown to be associated with malignancy in several types of cancers via formation of sialyl-Lewis<sup>x</sup> structures (attachment of an  $\alpha$ 1,3 linked-fucose to an antennary *N*-acetylglucosamine) which are linked to metastasis progression.<sup>142-144</sup> Furthermore,  $\alpha$ 2,6-linked sialic acids are presumed to be involved in blocking galectin binding and enhancing the tumor cell survival.<sup>8,9,145</sup>

Glycosylation, and especially changes thereof, has become of great importance due to the potential use of glycopeptides as clinical biomarkers.<sup>91,97,146</sup> Similarly, the structural characterization of protein-linked glycans is essential in the biopharmaceutical industry to ensure the safety and efficacy of novel therapeutics.<sup>133,134,147,148</sup> In recent years, molecular engineering of carbohydrates, referred to as glyco-engineering, has emerged in this field to enhance efficacy and design functional properties of monoclonal antibodies (mAbs).<sup>149-151</sup> Notably, terminal carbohydrate moieties on glycan chains are often affected by subtle changes in the cell culture environment.<sup>147,148,152</sup>

The characterization of protein glycosylation often requires information about site occupancy as well as about the glycan structures. This information is often obtained via proteolytic cleavage and subsequent glycopeptide analysis by mass spectrometry (MS).<sup>127,153</sup> However, an exhaustive characterization of sialic acid linkages of glycopeptides using MS remains rather challenging. Only recently, ion mobility mass spectrometry has demonstrated to allow for the separation of glycopeptide sialic acid linkage isomers.<sup>154</sup> Likewise, the differentiation between the differently linked sialic acids of immunoglobulin G (IgG) fragment crystallizable



(Fc) glycopeptides has been achieved using a linkage-specific derivatization method prior to MS analysis.<sup>155</sup> Nonetheless, such methods are not broadly accessible and there is a clear need for the differentiation of sialic acid isomers in conjunction with MS detection.

In this study, we present a novel approach based on capillary electrophoresis (CE) - electrospray ionization (ESI)-MS for glycopeptides analysis, enabling baseline separation of the  $\alpha$ 2,3- and  $\alpha$ 2,6-sialylated isomers without the need for any sample pre-treatment. The CE-ESI-MS method was developed using IgG Fc glycopeptides derived from recombinant IgG which contains only  $\alpha$ 2,3-sialylated glycopeptides,<sup>147,156</sup> as well as IgG Fc glycopeptides derived from human plasma IgG, which only contains  $\alpha$ 2,6-sialylated glycopeptides.<sup>157</sup> The observed difference in electrophoretic mobilities was further investigated by determining the difference in  $pK_a$  values of two representative compounds, namely,  $\alpha$ 2,3-sialyllactose and  $\alpha$ 2,6-sialyllactose. The relative difference in acidity constant was assessed using a novel methodology based on the internal standard (IS)-CE approach first proposed by Fuguet *et al.*,<sup>158</sup> identifying differences in acidity between sialic acid isomers as one of the plausible causes for sialic acid isomer separation. Finally, the developed platform was applied to discriminate sialic acid linkage isomers of tryptic glycopeptides of prostate-specific antigen (PSA).<sup>34</sup>

## MATERIALS AND METHODS

### MATERIALS

Ammonium bicarbonate (ABC), methanol (MeOH) of LC-MS grade and sodium dodecyl sulfate (SDS) were obtained from Merck (Darmstadt, Germany). Glacial acetic acid (HAc), DL-dithiothreitol (DTT), iodoacetamide (IAA), 1-hydroxybenzotriazole (HOBt) hydrate and Nonidet P-40 (NP-40) were acquired from Sigma-Aldrich (Steinheim, Germany). Ammonium acetate (AAC), formic acid (FA), trifluoroacetic acid (TFA) and water of LC-MS grade were purchased from Fluka (Steinheim, Germany). Sequencing Grade Modified Porcine Trypsin was obtained from Promega (Madison, WI). Phosphoric acid ( $H_3PO_4$ ), MES (2-(N-morpholino)ethanesulfonic acid), MOPS (3-morpholinopropanesulfonic acid), TRICINE (N-(tris(hydroxymethyl)-methyl)glycine), and CHES (2-(cyclohexylamino)ethane sulfonic acid) were purchased from Fluka (Buchs, Switzerland). Milli-Q water (MQ) was used for the matrix-assisted laser desorption ionization time-of-flight MS (MALDI-TOF-MS) experiments and obtained using a Q-Gard 2 system (Millipore, Amsterdam, Netherlands). HPLC supraGradient acetonitrile (MeCN) originated from Biosolve (Valkenswaard, Netherlands). 1-Ethyl-3-(3-(dimethylamino)propyl)-carbodiimide (EDC) hydrochloride was purchased from Fluorochem (Hadfield, United Kingdom). For the  $pK_a$  determination, HPLC grade MeOH was supplied by Romil (Kölliken, Switzerland), analytical reagent grade acetone by Acros Organics (Basel,



Switzerland), hydrochloric acid by Riedel-de-Haën (Buchs, Switzerland), and 1 M TitriNorm sodium hydroxide by VWR (Dietikon, Switzerland). The CEofix® pH 2.5 solutions were obtained from Analis (Namur, Belgium). Recombinant peptide-*N*-glycosidase F (PNGase F) was acquired from Roche Diagnostics (Mannheim, Germany) and 2,5-dihydroxybenzoic acid (2,5-DHB) from Bruker Daltonics (Bremen, Germany). Lyophilized sialidase S (*Streptococcus pneumoniae*) was obtained from PROzyme (Hayward, CA) and reconstituted in 20 µL MQ prior to use. Recombinant monoclonal IgG1 antibody (IgGmAb1) is derived from Chinese hamster ovary (CHO) cells. Human polyclonal IgG (intravenous immunoglobulin; IVIgG) was provided by Sanquin Research (Amsterdam, The Netherlands). Prostate-specific antigen (PSA) was acquired from Lee BioSolutions (St. Louis, MO). The two oligosaccharide standards 3'-sialyllactose sodium salt and 6'-sialyllactose sodium salt were obtained from Carbosynth (Compton, United Kingdom).

### TRYPTIC DIGEST

A PSA solution (1 µg/µL) was prepared in 25 mM ABC buffer. The reduction of disulphide bridges was carried out at 60°C for 30 min after addition of DTT (final concentration 2 mM) to the PSA solution. After cooling at room temperature, the sulfide alkylation was performed using IAA (final concentration 6 mM) and the sample was kept in the dark for 30 min. Sulfide alkylation reaction was stopped by addition of DTT (final concentration 6 mM). Finally, porcine trypsin was added to the sample (1:30, w/w) followed by overnight digestion at 37°C. Tryptic digests of IgGmAb1 and IVIgG solutions were prepared in 25 mM ABC buffer (1 µg/µL) with the addition of porcine trypsin (Promega Madison, WI, enzyme:sample 1:30, w/w). Digestion was performed overnight at 37°C.

### N-GLYCAN RELEASE

*N*-glycans were released from PSA as described previously.<sup>159</sup> Briefly, 10 µL of 2% SDS was added to 5 µL of PSA (21.5 µg) for denaturation and incubated for 10 min at 60°C. A mixture (10 µL) consisting out of 4% NP-40, 5x PBS and 1 mU PNGase F (10:10:1, v/v/v) was added for the release of the *N*-glycans and incubated overnight at 37°C.

### EXOGLYCOSIDASE TREATMENT

Ten microliters of the digested PSA sample (5.67 µg) was lyophilized and reconstituted in a mixture of 1 µL enzyme sialidase S, 1 µL of 50 mM sodium acetate (pH 5.5) and 8 µL of MQ, followed by overnight incubation at 37°C.

### ETHYL ESTERIFICATION

Ethyl esterification was performed as described by Reiding *et al.*<sup>160</sup> Briefly, the ethylation reagent was a mixture of EDC with HOBt dissolved in ethanol with a final concentration at 0.25 M for both compounds. One microliter of the PNGase F treated PSA was added to 20 µL of ethylation reagent and incubated for 1h at 37°C.

## GLYCAN ENRICHMENT BY COTTON HILIC SPE

Glycan enrichment was performed using cotton wool hydrophilic interaction liquid chromatography (HILIC) solid phase extraction (SPE) microtips, with a protocol modified from Selman *et al.*<sup>161</sup> Pipette tips with a volume of 20  $\mu\text{L}$  (Rainin Instrument, Oakland) were filled with approximately 3-mm long cotton thread (Pipooos, Netherlands). This HILIC-SPE sorbent was activated with 20  $\mu\text{L}$  MQ prior to conditioning with a solution of 85% MeCN (*v/v*), each step being repeated three times. Approximately 0.9  $\mu\text{g}$  of the ethylated PSA glycans was dissolved in 85% MeCN and loaded onto the cotton thread. The ethylated glycans were washed, first with a 20  $\mu\text{L}$  solution of MeCN-MQ-TFA, 85:14:1 (*v/v/v*) followed by 20  $\mu\text{L}$  with a solution of 85% MeCN, both steps were repeated three times. Analytes were eluted with 20  $\mu\text{L}$  of MQ.

## DETERMINATION OF $\text{pK}_\text{A}$ VALUES

Details about the theory and comprehensive explanations on the experimental procedure are available in **Section S-2-2** and **S-2-3** of the **Supporting information**.

In order to investigate the possible difference in  $\text{pK}_\text{a}$  values for  $\alpha$ 2,3- and  $\alpha$ 2,6-sialyllactose, an *in silico* prediction was initially performed as described in **S-2-3-2, Supporting information**.

$\text{pK}_\text{a}$  values were then experimentally determined using CE with ultraviolet (UV) detection. CE-UV experiments were performed with an HP<sup>3D</sup>CE system (Agilent Technologies, Waldbronn, Germany) equipped with an on-capillary diode array detector (8.5 cm from the anode), an autosampler, and a power supply able to deliver up to 30 kV. UV detection was carried out at 195 nm for target compounds and 260 nm for acetone (neutral marker of electro-osmotic flow, EOF). CE ChemStation (Agilent Technologies) was used for CE and UV control, data acquisition, and data handling. Analyses were performed with fused silica capillaries (BGB Analytik AG, Böckten, Switzerland) with 50  $\mu\text{m}$  i.d., 32.5 cm total length, and 8.5 cm effective length. The capillaries were coated with a dynamic coating (CEofix<sup>®</sup>, Analis, Namur, Belgium) to generate a large EOF (*ca.*  $> 2.2 \cdot 10^{-4} \text{ cm}^2 \cdot \text{V}^{-1} \cdot \text{s}^{-1}$ ) regardless of the buffer pH. Samples were kept at room temperature in the autosampler and introduced by short-end injection, namely, the polarity was reversed and samples were injected at the detector side. Applying a pressure of 12 mbar for 4 s performed an injection equivalent to 0.4% of the capillary total length. The capillary was thermostated at 25°C by a high velocity air stream. Before its first use, the capillary was sequentially rinsed (1 bar) with MeOH, 1 M hydrochloric acid (HCl), water, 1 M sodium hydroxide (NaOH), 0.1 M NaOH, and water (5 min each) prior to the coating procedure. The latter involved capillary rinses with aqueous polymer solutions (CEofix<sup>®</sup> initiator and accelerator). Between different pH buffers, several washing steps (1 bar) were employed, namely water (1 min), CEofix<sup>®</sup> initiator (0.5 min), CEofix<sup>®</sup> accelerator (1 min), and background electrolyte (BGE; 5 min), followed by pre-electrophoresis (3.5-8 kV, 5 min). Between analyses, the capillary was rinsed (1 bar) with buffer for 1 min. When not in use, the capillary was rinsed with water and stored dry.



For the determination of absolute  $pK_a$  values, a voltage of 3.5 to 8 kV (depending on the running buffer, see **Table S-2.2, Supporting information**) was applied to avoid detrimental effect of Joule heating. Other experimental conditions for the absolute  $pK_a$  values are described in **S-2-3-3, Supporting information**.

For the relative  $pK_a$  determination, two phosphate buffers at pH 2.0 and 6.0 were prepared. They were set at a constant ionic strength of 50 mM to avoid variations in effective mobility and keep activity coefficients constant (**Table S-2.3, Supporting information**). The buffer pH values were selected to ensure that both forms coexist (*i.e.*, at pH 2.0) and only the fully deprotonated form exists (*i.e.*, at pH 6.0). The pH values were measured with a Mettler-Toledo SevenMulti pH meter that was calibrated on a daily basis. The injected sample was a mixture of both investigated compounds at a concentration of 0.1 mg/mL in MQ and acetone 95:5 (*v/v*). They were injected twice ( $N = 2$ ) at each pH. Voltages of 8 kV and 5 kV were applied when using the BGEs at pH 6.0 and 2.0, respectively.

## MALDI-TOF-MS

MS experiments of ethylated PNGase F released *N*-glycans were performed with MALDI-TOF-MS (UltrafleXtreme, Bruker Daltonics, Bremen, Germany) using FlexControl 3.3 software (Bruker Daltonics) in positive ionization mode. The ethylated PNGase F released *N*-glycans were spotted on an 800  $\mu\text{m}$  AnchorChip plate (Bruker Daltonics) together with 1  $\mu\text{L}$  of 2,5-DHB matrix (5 mg/mL, 1 mM NaOH in MeCN-MQ, 50:50, *v/v*). A peptide calibration standard (Bruker Daltonics) was used for external calibration. FlexAnalysis (Bruker Daltonics) was used for the analysis of the acquired MALDI-TOF-MS data. Modifications of the sialic acids were taken into account (lactonization or ethylation) and all glycan masses were recorded using  $[M + \text{Na}]^+$  ions.

## CAPILLARY ELECTROPHORESIS - ELECTROSPRAY IONIZATION - MASS SPECTROMETRY

All CE experiments were carried out on a CESI 8000 system (SCIEX Separations, Framingham, MA), and 90-cm long bare fused capillaries (i.d. of 30  $\mu\text{m}$ , o.d. of 150  $\mu\text{m}$ ) were used. Prior to analysis, the capillary was rinsed thoroughly with 0.1 M NaOH (2.5 min), 0.1 M HCl (2.5 min), water (4 min) and BGE which consisted of 10% HAc (*v/v*, pH 2.3). The sample was on-line pre-concentrated using transient isotachopheresis (t-ITP). Prior to injection, AAC at pH 4.0 (final concentration 100 mM) was added to the sample, which acted as the leading electrolyte. Samples were injected using hydrodynamic injection of 1 psi for 60 s (9 nL, corresponding to 1.4% of the capillary volume). For all experiments, a BGE post plug was added by applying 0.5 psi for 25 s (corresponding to 0.3% of the capillary volume). Separation was carried out by applying a voltage of 20 kV.

The CE system was hyphenated to a UHR-QqTOF maXis Impact HD MS (Bruker Daltonics) via a sheathless CE-ESI-MS interface from SCIEX. A capillary voltage of -1300 V was applied to ensure a stable electrospray. All experiments were performed in positive ionization mode. The drying gas (nitrogen) flow rate and temperature were set at 1.5 L/min and 180°C, respectively. MS data was acquired between  $m/z$  200-2200 with a spectral acquisition rate of 1 Hz.

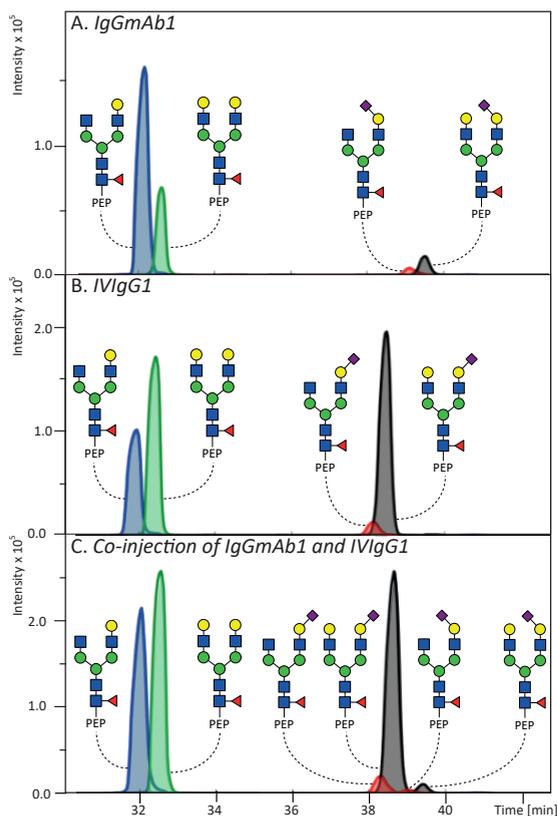
CE-ESI-MS data was analyzed with DataAnalysis 4.2 (Build 387, Bruker Daltonics). Prior to data analysis, all MS spectra were calibrated using sodium adducts detected at the beginning of the electropherogram. The data was manually screened for glycopeptides and compositions were matched based on the exact mass, selectivity and relative intensities (**Table S-2.1, Supporting information**). Fragmentation spectra were acquired for 27% of the identified compositional glycopeptide variants of PSA (**Figure S-2.4 - S-2.7, Supporting information**). Extracted ion electropherograms (EIEs, smoothed with a Gaussian fit) were acquired with the first three isotopes of the doubly, triply and quaternary charged analytes using a width of  $\pm 0.05$   $m/z$  unit.

For data alignment, to ensure comparable migration times and reveal the separation of sialic acid linkage isomers, the analysis of IgGmAb1 and IVIgG glycopeptides used 2-aminobenzamide labeled dextran ladder as an internal calibrant. For the alignment of PSA glycopeptides no IS was needed as the sample was rich enough to ensure proper alignment without an IS. The alignment program is an in-house and open source program, based on the LaCyTools alignment tool.<sup>162</sup> The alignment program is supplied with this manuscript and can be found in **S-2-4** of the **Supporting information**. Briefly, an alignment procedure referred to as “targeted alignment” was achieved by supplying a list (.txt file) of the calibrants with their theoretical  $m/z$  and expected migration time (mt) values. After converting the data file to an mzML file, the algorithm examines the mt and  $m/z$  regions containing the calibrants. The local maximum within each region is isolated. The mt matching to the local maximum is then coupled to the specified mt of the calibrant. All observed mt and expected mt combinations are then taken as an input array for a second-degree polynomial fit. The function ( $f$ ) returned by the polynomial fit is used to transform the observed mt of the whole spectrum by performing  $mt(new) = f(mt(old))$ , resulting in well-aligned spectra. This alignment strategy is opposed to an unbiased alignment,<sup>163</sup> which uses random selection of features and is typically suited for unknown samples that contain a wide range of analytes. Targeted alignment is more adequate for highly purified samples in which the analytes are well known and characterized.<sup>162</sup> Afterwards, the aligned runs were plotted using an in-house developed and open source plotting program (**S-2-5** of the **Supporting information**).

## RESULTS AND DISCUSSION

### ANALYSIS OF IgG FC GLYCOPEPTIDES USING CE-ESI-MS

A sheathless CE-ESI-MS platform was used for the analysis of Fc glycopeptides derived from IVIgG and IgGmAb produced in CHO cells. In order to correct for the significant differences observed in mt, a targeted alignment of the data was applied using an IS. **Figure 2.1** illustrates the analysis of IgGmAb1 (**Figure 2.1.A**) and IVIgG (**Figure 2.1.B**) after alignment. For both samples, the extracted ion electropherograms of the non-sialylated glycopeptides of subclass 1 (G1F and G2F) showed the same apparent mobility while different mt were obtained for the mono-sialylated species (G1FS and G2FS). IgG Fc glycopeptides derived from CHO cells are well-known to contain only  $\alpha$ 2,3-linked sialic acids while IgG Fc glycopeptides derived from human plasma are expected to only contain  $\alpha$ 2,6-linkages.<sup>147,156,157</sup> The co-injection of IgGmAb1 and IVIgG1 glycopeptides (**Figure 2.1.C**) emphasized this difference in apparent mobilities for  $\alpha$ 2,3- versus  $\alpha$ 2,6-linked sialylated species, where  $\alpha$ 2,3-linked mono-sialylated and  $\alpha$ 2,6-linked mono-sialylated glycopeptides were baseline resolved. The differentiation between both linkages is highly relevant since  $\alpha$ 2,3- and  $\alpha$ 2,6-linkages may have significantly different biological effects.<sup>155</sup> CE is not the only separation platform that



**Figure 2.1: Extracted ion electropherograms (EIEs) of IgGmAb1 and IVIgG1 glycopeptides obtained with CE-ESI-MS after targeted alignment. (A)** EIEs of IgGmAb1 glycopeptides derived from CHO cells, **(B)** EIEs of IVIgG1 retrieved from human plasma and **(C)** EIEs of a co-injection of IgGmAb1 and IVIgG1. The “PEP” label illustrates the tryptic peptide sequence EEQYNSTYR to which the glycan is attached.

enables separation of sialylated isomers: zwitterionic -HILIC and porous graphitized carbon (PGC)-LC have proven to be suitable for separation of glycan species with isomeric sialic acid linkages.<sup>164,165</sup> Specific fragmentation features have been reported for anionic glycans, enabling the distinction between isomers.<sup>166,167</sup> However, many other conventional positive-mode MS(/MS) approaches do not provide sufficient differentiation between  $\alpha$ 2,3- and  $\alpha$ 2,6-linked sialylated species due to an identical exact mass as well as similar fragmentation patterns using collision induced dissociation.<sup>168</sup>

Our preliminary data obtained for the analysis of IgG Fc glycopeptides using CE-ESI-MS showed different electrophoretic mobilities of sialic acid linkage isomers, implying a different charge-to-size ratio that may be used to discriminate both  $\alpha$ 2,3- and  $\alpha$ 2,6-linked sialylated species. This difference may be either caused by differences in the hydrodynamic radius of the analytes and/or by a difference in acidic constant between  $\alpha$ 2,3- and  $\alpha$ 2,6-linked sialic acids. To the best of our knowledge, only a general  $pK_a$  value for both linkages ( $pK_a$  ca. 2.6-2.9) has been described so far.<sup>169-171</sup> Therefore, the potential difference in  $pK_a$  values between the two differently linked sialic acids was further investigated using two representative compounds, namely,  $\alpha$ 2,3-sialyllactose and  $\alpha$ 2,6-sialyllactose.

### **$pK_a$ DIFFERENCES BETWEEN $\alpha$ 2,3- AND $\alpha$ 2,6- SIALYLATED LACTOSE**

The *in silico* prediction as well as the experimental determination of the  $pK_a$  values are described in **Supporting information, Section S-2-1**.<sup>172</sup> As summarized in **Table 2.1**, both approaches were not sufficient to demonstrate a significant difference in the  $pK_a$  values of  $\alpha$ 2,3- and  $\alpha$ 2,6-linked sialyllactose due to the relative close values and the associated error with the measurements. In order to further explore the expected difference in  $pK_a$  values and to enable a clear differentiation between both  $pK_a$  with minimized measurement errors, a novel methodology based on the so-called IS-CE approach was implemented for the determination of the relative difference in  $pK_a$  values. This IS-CE approach, first proposed by Fuguet *et al.*,<sup>158</sup> is based on the assessment of acidity constants using CE with concomitant injection of the analyte of interest and an IS. The  $pK_a$  values of the latter are known and its selection is based on its close similarity in nature and  $pK_a$  value to the analyte of interest. Here, this procedure was adapted to assess the relative difference in  $pK_a$  unit where one of the sialyllactose acted as the IS and the other one as the analyte of interest. The relative difference was estimated by calculating the mean limiting mobilities (migration under fully deprotonated form, namely, at pH 6.0,  $N = 2$ ) and mean effective

**Table 2.1:  $pK_a$  values determined for analogs of sialyllactose.**

ACD/Lab method	Predicted $pK_a$ value
$\alpha$ 2,3 sialyllactose	1.9 ( $\pm$ 0.7)
$\alpha$ 2,6 sialyllactose	2.0 ( $\pm$ 0.7)
Conventional CE method	Absolute $pK_a$ value
$\alpha$ 2,3 sialyllactose	1.8 ( $\pm$ 0.1)
$\alpha$ 2,6 sialyllactose	1.8 ( $\pm$ 0.1)
IS-CE method (N = 2)	Relative $pK_a$ value
	$3.4 \cdot 10^{-2}$

mobilities (migration under partially ionized form, namely, at pH 2.0,  $N = 2$ ) for both analytes ( $\alpha$ 2,3- and  $\alpha$ 2,6-linked sialyllactose) injected simultaneously therefore no alignment was needed (**Figure S-2.1, Supporting information**). Mean limiting mobilities for both compounds were found to be  $-6.0 \cdot 10^{-5} \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$  ( $\pm 3.4 \cdot 10^{-6} \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ ) and  $-6.5 \cdot 10^{-5} \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$  ( $\pm 3.4 \cdot 10^{-6} \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ ), while the mean effective mobilities were estimated at  $-4.2 \cdot 10^{-5} \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$  ( $\pm 9.7 \cdot 10^{-7} \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ ) and  $-4.6 \cdot 10^{-5} \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$  ( $\pm 8.6 \cdot 10^{-7} \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ ). Using the measured mobilities and **Eqs. S-2-7, S-2-8 and S-2-9 in Section S-2-2 of the Supporting information**, a  $\Delta pK_a$  of  $3.4 \cdot 10^{-2}$  unit was assessed (**Table 2.1**).

This relative  $\Delta pK_a$  therefore reflects a difference in the acidity constants of the two sialylated glycan isomers. Consequently, this led to different electrophoretic mobilities in the working BGE (*i.e.*, 10% HAc at pH 2.3) for  $\alpha$ 2,3- and  $\alpha$ 2,6-linked sialylated glycopeptides of IgGmAb1 and IVIgG. Considering that other experimental conditions are used in the CE-ESI-MS approach compared to the IS-CE approach (*i.e.*, longer capillary (90 cm *versus* 32.5 cm), long-end injection *versus* short-end injection (89 cm *versus* 8.5 cm effective length of the capillary) and a higher voltage (5/8 kV *versus* 20 kV)), a migration time difference of 1 min with MS detection was estimated between the  $\alpha$ 2,3- and  $\alpha$ 2,6-linked sialyllactose on the basis of the deduced  $pK_a$  difference (see **Eq. S-2.7 in Section S-2.2 of the Supporting information**). This finding supports the result obtained with the co-injection of IgGmAb1 and IVIgG1 illustrated in **Figure 2.1**. The different linkage affects the charge of the sialyllactoses, which may also present a different size, overall impacting the charge-to-size ratio and, thus, the migration behavior. Therefore, the CE-ESI-MS technique appears to be a promising tool for separating sialic acid linkage isomers.

## ANALYSIS OF PROSTATE-SPECIFIC ANTIGEN GLYCOPEPTIDES USING CE-ESI-MS

Due to the heterogeneous glycosylation character of PSA, including differentially linked sialic acids, this glycoprotein was selected to investigate the potential of using of CE-ESI-MS for sialic acid linkage isomer differentiation.<sup>3</sup> A state-of-the-art MALDI-TOF-MS platform was also used as orthogonal technique after enzymatic release of the glycans from PSA and an ethyl esterification to distinguish  $\alpha$ 2,3- from  $\alpha$ 2,6-linked sialic acids of released glycans, the results were compared to the CE-ESI-MS platform.<sup>160</sup> Moreover, in order to evaluate the migration behavior of the sialylated isomers of PSA, an exosialidase step was performed to allow for the release of  $\alpha$ 2,3-linked sialic acids from complex glycopeptides.

### ETHYL ESTERIFICATION

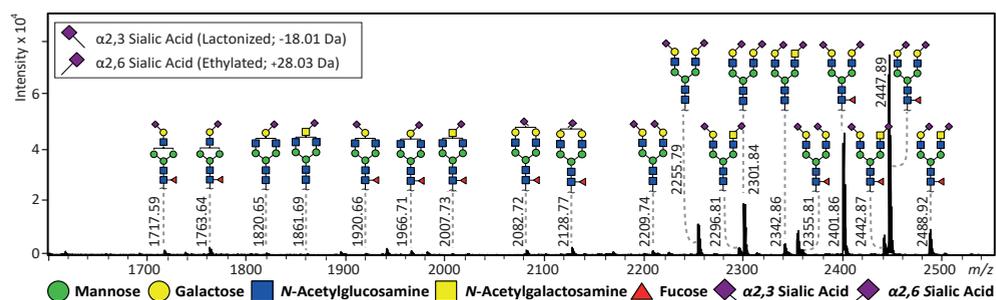
The enzymatically released *N*-glycans from PSA were derivatized using a procedure developed by Reiding *et al.*, converting  $\alpha$ 2,3-linked sialic acids into a lactone (- 18 Da) while  $\alpha$ 2,6-linked sialic acids formed an ethyl ester (+ 28 Da).<sup>160</sup> The MALDI-TOF-MS profiling confirmed that



the acquired PSA consisted of differentially linked sialic acids (**Figure 2.2**), enabling the identification of 37 glycans (**Table S-2.1, Supporting information**). It is worth mentioning that sialic acid attached to a *N*-acetylgalactosamine only revealed an  $\alpha$ 2,6-linkage, which is supported by literature.<sup>82</sup>

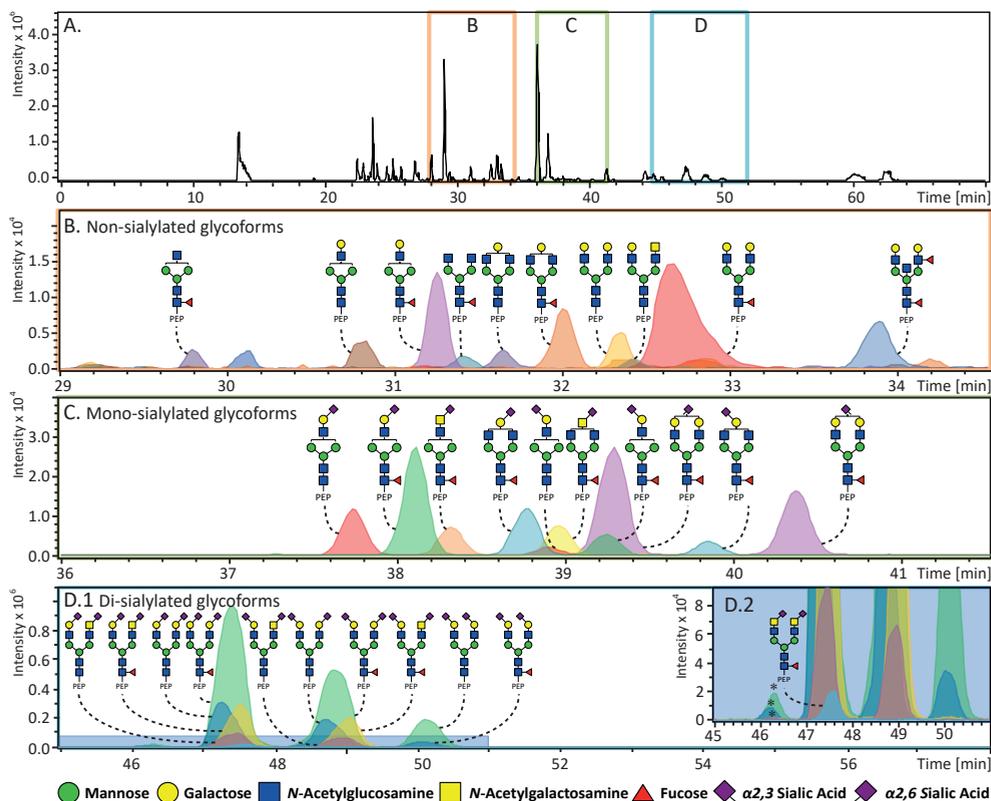
### CE-ESI-MS(/MS) ANALYSIS OF A PSA TRYPTIC DIGEST

**Figure 2.3** displays the base peak electropherogram and the EIEs obtained for the analysis of a PSA tryptic digest. A relatively high number of glycopeptides (*i.e.*, 75 glycopeptides,  $S/N > 9$ , ppm error  $< 10$ ) were identified with the CE-ESI-MS setup. All glycans were attached to the peptide “N<sub>69</sub>K” and are listed in **Section S-2-6 of the Supporting information, Table S-2.1**. Notably, a relatively low amount of PSA (*ca.* 1 ng protein weight injected) was necessary to detect these 75 glycopeptides. The base peak electropherogram of a PSA tryptic digest analysis is shown in **Figure 2.3.A** while the EIEs of non-sialylated, mono-sialylated, and di-sialylated glycoforms are shown in **Figure 2.3.B-D**. **Figure 2.3.C** and **2.3.D** clearly emphasize that  $\alpha$ 2,3- and  $\alpha$ 2,6-linked isomers are baseline resolved, confirming the ability of CE-ESI-MS to discriminate between the different linkages, independently of the glycopeptide structure. The identified compounds are represented in terms of number of hexoses (H), *N*-acetylhexosamines (N), fucoses (F) and sialic acids (S). As illustrated in **Figure S-2.2** and **S-2.3** of the **Supporting information**, isomer separation was observed for the  $\alpha$ 2,3- and  $\alpha$ 2,6-linked sialylated species (H5N4S2, H5N4F1S2, H4N5S2 and H4N5F1S2). Interestingly, a peak with a higher apparent mobility compared to the fully  $\alpha$ 2,6-di-sialylated species was detected, with a similar isotopic distribution. This peak could not be assigned with MS/MS due to the relatively poor signal intensity. Given the migration in the di-sialylated region of the electropherogram, this unidentified analyte may also be a glycopeptide containing two sialic acids, from which one is an  $\alpha$ 2,8-linked sialic acid on top of the second sialic acid. However, the presence of a di-sialic acid motif with an  $\alpha$ 2,8-linkage could not be confirmed with the MALDI-TOF experiments as this  $\alpha$ 2,8-linked sialic acid is expected to readily form lactones under ethyl esterification conditions, leading to the same  $m/z$  value as for a glycan with an



**Figure 2.2:** MALDI-TOF-MS spectrum of PNGase F released PSA N-glycans treated with ethyl esterification. In total, 37 different glycans were identified (not all data shown, a complete overview can be found in **Table S-2.3, Supporting Information**).

$\alpha$ 2,3-linked sialic acid even though there is a difference in linkage position.<sup>173</sup> Next to that, there are hitherto no IS available that would contain this specific linkage. Therefore, similar experiments as performed with the  $\alpha$ 2,3- and  $\alpha$ 2,6-linked sialyllactose with IS-CE could not be performed. As a result, further investigations are needed to confirm this hypothesis.



**Figure 2.3:** CE-ESI-MS analysis of PSA tryptic (glyco)peptides. **(A)** Representative base peak electropherogram observed for a tryptic digest of PSA. Based on the electrophoretic separation, three distinct clusters were defined. **(B)** Cluster with only non-sialylated glycopeptides of PSA. **(C)** Cluster containing mono-sialylated glycopeptides of PSA. **(D)** Cluster with di-sialylated glycopeptides of PSA. In total, 75 different glycopeptides were identified (not all data shown, a complete overview can be found in **Table S-2.3, Supporting Information**). Peaks with a star (\*) were not assigned. The “PEP” label illustrates the tryptic peptide sequence N<sub>69</sub>K to which the glycan is attached.

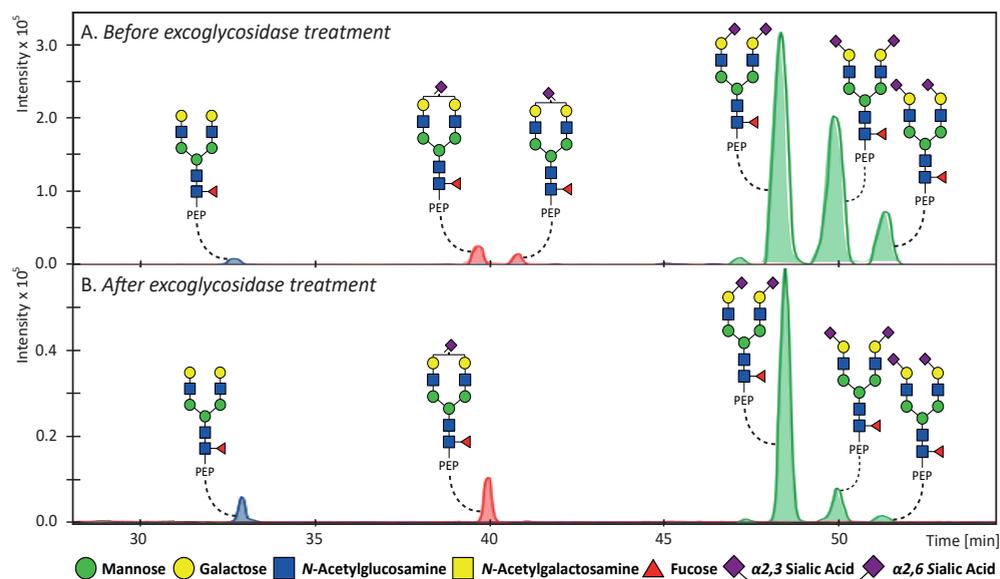
Next to the isomer separation of the differentially linked sialic acids, **Figure 2.3** also highlights that the glycopeptide separation is highly influenced by the degree of sialylation. This is expected as the migration of the glycopeptide is dependent on the charge-to-size ratio of the analyte in the capillary. **Figure 2.3.B** illustrates that the non-sialylated glycopeptides migrate first out of the capillary followed by the mono- and di-sialylated species. The sialic acids are expected to carry a negative charge when 10% HAC is used as BGE (pH 2.3), resulting in a slower migration in the capillary. More sialic acids on the analyte will therefore increase the



migration time. **Figure 2.3.B** illustrates the EIEs observed for non-sialylated glycopeptides and revealed no isomer separation (co-migration of isomers), indicating that the separation mostly relies on the number of sialic acids attached to the analyte as well as the  $\Delta pK_a$  of the differentially linked sialic acids. Differences in hydrodynamic radii of the analytes appeared to have only a minor influence, as illustrated for glycopeptides with neutral glycans containing zero, one or two galactose residues, which only showed minor differences in migration without differentiation between isomers.

### CE-ESI-MS ANALYSIS AFTER EXOSIALIDASE TREATMENT

**Figure 2.4** shows the EIEs observed for the di-sialylated glycopeptides H5N4F1S2 (blue trace) and H4N5F1S2 (green trace) prior (**Figure 2.4.A**) and after (**Figure 2.4.B**) treatment with  $\alpha$ 2,3-sialidase. Both analyses were aligned using the targeted alignment program to ensure that a proper identification could be made of the different isomeric species between the two analysis. The electropherograms obtained for  $\alpha$ 2,3- and  $\alpha$ 2,6-sialylated species indicated a significant decrease of the relative abundance of  $\alpha$ 2,3-sialylated variants. The intensities observed for  $\alpha$ 2,6-linked mono-sialylated glycopeptide (H5N4F1S1<sub>2,6</sub>) and the non-sialylated glycopeptide (H5N4F1) increased compared to the most abundant peak ( $\alpha$ 2,6-linked di-sialylated glycopeptide, H5N4F1S2<sub>2,6</sub>) while all  $\alpha$ 2,3-linked sialylated species (H5N4F1S1<sub>2,3</sub>, H5N4F1S1<sub>2,3</sub>S<sub>2,6</sub> and H5N4F1S2<sub>2,3</sub>) showed a decreased signal. However, an incomplete



**Figure 2.4:** CE-ESI-MS analysis of PSA with and without  $\alpha$ 2,3-sialidase treatment using CE-ESI-MS after targeted alignment. **(A)** EIEs of the non-sialylated glycopeptide (H5N4F1), the mono-sialylated glycopeptides (H5N4F1S1) and the di-sialylated glycopeptides (H5N4F1S2) of PSA prior to exoglycosidase treatment and **(B)** after exoglycosidase treatment.

reaction was observed for the exoglycosidase treatment, which may most likely be attributed to experimental conditions such as (too short) incubation time and (low) amount of enzyme added. Nonetheless, these results confirmed that the CE-ESI-MS setup enabled a baseline separation of  $\alpha$ 2,3- and  $\alpha$ 2,6-sialylated glycopeptides, with  $\alpha$ 2,3-variants showing a slower migration.

## SYNOPSIS

Little attention has been devoted to the analysis of glycopeptides with CE-ESI-MS so far. This is unfortunate since this technique has shown high separation efficiency and sensitivity in the context of large biomolecules analysis, especially in combination with a sheathless ESI interface.<sup>112,174-176</sup> For instance, Heemskerck *et al.*<sup>177</sup> showed that a 40-fold increase in sensitivity with CE-ESI-MS compared to nano-LC-ESI-MS was observed for the analysis of plasma derived IgG1 Fc glycopeptides.

This study presented a novel approach for in-depth and direct analysis of complex glycoproteins with differently linked sialic acids using a CE-ESI-MS platform, without the need for tedious and time-consuming sample preparation.<sup>34,155,160</sup> With the developed setup, 75 different glycopeptides with one single *N*-linked glycosylation site, position R-N<sub>69</sub>-K-S, were detected for the analysis of PSA tryptic digest. This CE-ESI-MS platform shows a great potential for PSA glycopeptide analysis with results comparable to an interlaboratory study performed on PSA and PSA high isoelectric point isoform (bulk-purified proteins based on isoelectric point).<sup>178</sup> The latter reported that 61 *N*-glycan compositions could be detected using multiple analytical strategies including bottom-up *versus* top-down analysis, released *N*-glycans analysis by PNGase F, polymeric reversed-phase *versus* monolith and C8 chromatography, as well as CE.

Nowadays, the concentration of PSA in serum is widely used in clinics for the screening of prostate cancer (PCa) since elevated PSA serum levels can be indicative of malignancy.<sup>179</sup> However, this screening method shows a poor specificity as various other causes also lead to increased PSA levels in serum.<sup>180,181</sup> Previous studies indicated that alterations in the glycosylation profile of PSA could provide a more specific tool for diagnosis and prognosis of PCa.<sup>3,46,99,182-184</sup> This study presents a compelling platform for the analysis of these alterations, offering a very competitive tool to improve both specificity and sensitivity. Moreover, it allows to evaluate the PSA glycosylation, a promising approach for early detection, for better prognosis or staging of PCa. However, while currently one ng of protein weight was injected and analyzed, such amounts will be difficult to retrieve from serum of PCa patients, who often have PSA concentrations lower than 10 ng/mL. On the opposite, urine PSA concentrations of PCa patients have been reported with a median value of 425 ng/mL, a sufficient concentration level for this CE-ESI-MS platform.<sup>5</sup> Therefore, the potential of PSA biomarkers



derived from urine should be investigated and could open new possibilities for establishing a less invasive but hopefully more specific diagnostic marker than the current clinical PSA test. This has, to the best of our knowledge, not yet been explored. Future perspectives include the analysis of glycoproteins with higher complexity and more glycosylation sites, as well as highly sialylated glycopeptides which may show a negative apparent mobility, rendering their detection critical with the current setup. Further method development is therefore needed to achieve the analysis of, for instance, tri- or tetra-sialylated glycopeptides.

## CONCLUSION

This study presents the implementation of a CE-ESI-MS platform enabling the differentiation between  $\alpha$ 2,3- and  $\alpha$ 2,6-sialylated glycopeptides without prior derivatization. Where conventional MS(/MS) approaches cannot resolve these isomers due to identical molecular formulas and similar fragmentation patterns, CE enabled the baseline separation of sialylated glycopeptides due to a difference in their electrophoretic mobilities, correlated by a difference in acidity, as shown for the two representative compounds  $\alpha$ 2,3-sialyllactose and  $\alpha$ 2,6-sialyllactose. The developed platform was used for the analysis of tryptic PSA glycopeptides, enabling the identification of 75 PSA glycopeptides, a significantly higher amount compared to the 37 PNGase F released *N*-glycan signals detected by MALDI-TOF-MS. This highlights the gain in sensitivity obtained with the developed platform and its potential for biomarker discovery, enabling the possible differentiation between clinically relevant glycopeptides revealing differentially-linked sialic acids.

## ACKNOWLEDGEMENTS

Christophe Francey is acknowledged for his technical support regarding  $pK_a$  determination. We thank Dietmar Reusch (F. Hoffmann-La Roche AG), Theo Rispens (Sanquin Research), David Falck and Agnes Hipgrave Ederveen for providing samples. This research was supported by the European Union (Seventh Framework Programme HighGlycan and IBD-BIOM, Grant Agreements 278535 and 305479), a Horizon Programme Zenith project of The Netherlands Genomic Initiative (Project No. 93511033), Astellas Pharma B.V. and SCIEX. Isabelle Kohler was a recipient of a fellowship (P2GEP3\_155633) supported by the Swiss National Science Foundation.

## SUPPORTING INFORMATION

The full supplementary information of this thesis can be found via <https://doi.org/10.17026/dans-zg4-nxca>. In addition, the supporting information as referred to in G.S.M. Kammeijer et al., “Sialic Acid Linkage Differentiation of Glycopeptides using Capillary Electrophoresis - Electrospray Ionization - Mass Spectrometry”, *Scientific Reports* 7 (1), Article number: 3733, 2017, is available via <https://www.nature.com/articles/s41598-017-03838-y#Sec21>.



