



Article

# Site-Specific Glycan Microheterogeneity Evaluation of Aflibercept Fusion Protein by Glycopeptide-Based LC-MSMS Mapping

Ju Yeon Lee <sup>1</sup>, Jin-Woong Choi <sup>1,2</sup>, Seoyoung Hwang <sup>1,2</sup>, Sung Ho Hahm <sup>3</sup> and Yeong Hee Ahn <sup>4,\*</sup>

<sup>1</sup> Research Center for Bioconvergence Analysis, Korea Basic Science Institute, Cheongju 28119, Korea

<sup>2</sup> Graduate School of Analytical Science and Technology, Chungnam National University, Daejeon 34134, Korea

<sup>3</sup> Rophibio Inc., Cheongju 28160, Korea

<sup>4</sup> Department of Biomedical Science, Cheongju University, Cheongju 28160, Korea

\* Correspondence: ahnyh@cju.ac.kr; Tel.: +82-43-229-8565

**Abstract:** The evaluation of the protein glycosylation states of samples of aflibercept obtained from three different regions was conducted by site-specific N-linked glycan microheterogeneity profiling. Glycopeptide-based nano-LC MSMS mapping of tryptic-digested samples of each aflibercept lot provided site-specific information about glycan microheterogeneity on each of the five N-glycosylation sites (two sites in the VEGFR-1 region, two sites in the VEGFR-2 region, and one site in the human IgG Fc region). Next, the glycopeptide-mapping results obtained from the three different aflibercept lots were compared to evaluate the similarity between the samples. Three aflibercept lots showed a high degree of similarity in glycan composition, fucosylation level, sialylation level, and branching, when all five N-glycosylation sites were assessed together as a group. On the other hand, noticeable variations between lots in the glycan types and sialylation levels on the two sites of the VEGFR-2 domain were observed when each of the five N-glycosylation sites were assessed using the glycopeptide-based method. The presence of N-glycolylneuraminic acid (NeuGc) glycans, which may mediate adverse immune reactions in antibody therapeutics, were also detected on the sites of VEGFR1 and VEGFR2 domains, but not on the IgG Fc domain site. These results imply that analyses of the glycosylation profiles of fusion proteins containing multiple N-glycosylation sites, such as aflibercept, being done as a part of quality control for the therapeutics manufacturing process or for biosimilar development, can be done with a more applicable outcome by assessing each site separately.

**Keywords:** aflibercept fusion protein; glycan microheterogeneity; site-specific evaluation; glycopeptide-based mapping; LC-MSMS; N-glycolylneuraminic acid (NeuGc); I-GPA; q-GPA



**Citation:** Lee, J.Y.; Choi, J.-W.; Hwang, S.; Hahm, S.H.; Ahn, Y.H. Site-Specific Glycan Microheterogeneity Evaluation of Aflibercept Fusion Protein by Glycopeptide-Based LC-MSMS Mapping. *Int. J. Mol. Sci.* **2022**, *23*, 11807. <https://doi.org/10.3390/ijms231911807>

Academic Editor: Alexander O. Chizhov

Received: 22 August 2022

Accepted: 30 September 2022

Published: 5 October 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Protein glycosylation is a post-translational modification (PTM) that can impact a protein's function or biological activity, such as intracellular transport, cellular localization, secretion, protein-protein interaction, stability, or clearance. Consequently, protein glycosylation can become one of the most important quality attributes related to the development of protein therapeutics, including antibody or Fc-fusion protein therapeutics. Especially, changes in the efficacy or toxicity of protein therapeutics can occur as a result of changes in pharmacokinetics, pharmacodynamics or immunogenicity depending on the nature and extent of protein glycosylation [1–4]. Biosimilars are defined as biological products that are highly similar to the reference products, notwithstanding minor differences in clinically inactive components. Glycosylation profile is one of the most critical quality attributes (CQA), in terms of the safety and regulatory controls, that is required for biosimilars to closely match the reference products [5].

The majority of protein therapeutics, including aflibercept, are produced from animal cell-based manufacturing systems, after generating stable cell lines using animal cells, such as CHO cells as a host. In contrast to chemical synthesis processes that can be maintained consistently with precision between batches for the manufacturing of small molecule therapeutics, it is practically impossible to maintain identical conditions for the animal cell-based production systems, even when batches were manufactured at the same site with the same procedure. Variations can occur during the upstream bioreactor culturing process, in terms of pH, DO, and temperature, which will cause variations in cell growth, viability and cell density, which will then cause variations in the downstream purification process [6–9].

Since protein glycosylation is influenced by a variety of elements involved in the production conditions of the manufacturing system, the unique characteristics in glycan microheterogeneity of a therapeutic protein can obviously be varied during the production process. Therefore, a very robust quality control system will have to be applied to make sure that all the upstream and downstream processes are maintained as consistently as possible between batches of production, and to monitor and minimize variations in physicochemical attributes, such as glycosylation profile. The importance of the process and quality controls becomes more obvious for the development of biosimilars. In most cases, the developers of innovative therapeutics will keep their manufacturing process information as proprietary. Therefore, biosimilars developers will have to develop their own process with likely variations from that of the original developers, while maintaining similarity with the reference products with minimum variations in the critical quality attributes [10,11].

Mass spectrometry coupled with liquid chromatography is one of the most powerful tools that can be used to acquire high-throughput and accurate informatics for mapping protein glycosylation patterns [12,13]. Since the protein glycosylation pattern of antibody therapeutics is considered a critical quality attribute affecting a wide spectrum of the biological processes of the therapeutics, precise mapping of glycosylation patterns is very important to evaluate safety and efficacy of the drug [14,15]. Although 2-AB labeling for the glycans, released from IgG monoclonal antibodies, has been a useful glycan mapping method [16,17], this glycan-based mapping shows limitations for the analysis of proteins with multiple glycosylation sites by virtue of the loss of information on site-specific microheterogeneity of a multi-glycosylated proteins, like aflibercept fusion protein.

Aflibercept is a fusion protein between the human vascular endothelial growth factor receptors' extracellular domains and the Fc portion of human IgG1 [18]. It is comprised of vascular endothelial growth factor receptors 1 and 2 (VEGFR-1, VEGFR-2) fused to the Fc domain of human IgG1, and contains a total of five N-glycosylation sites (two sites from the VEGFR-1 region, two sites from the VEGFR-2 region, and one site from the human IgG Fc region). While most therapeutic antibodies contain a single N-glycosylation site on the Fc portion of the protein, aflibercept, which contains four extra N-glycosylation sites on its VEGF receptor domains, exhibits different glycosylation profiles among glycosylation sites [19]. Consequently, a method of analysis is required that is capable of differentiating the glycan microheterogeneity site-specifically of the multiply glycosylated fusion protein.

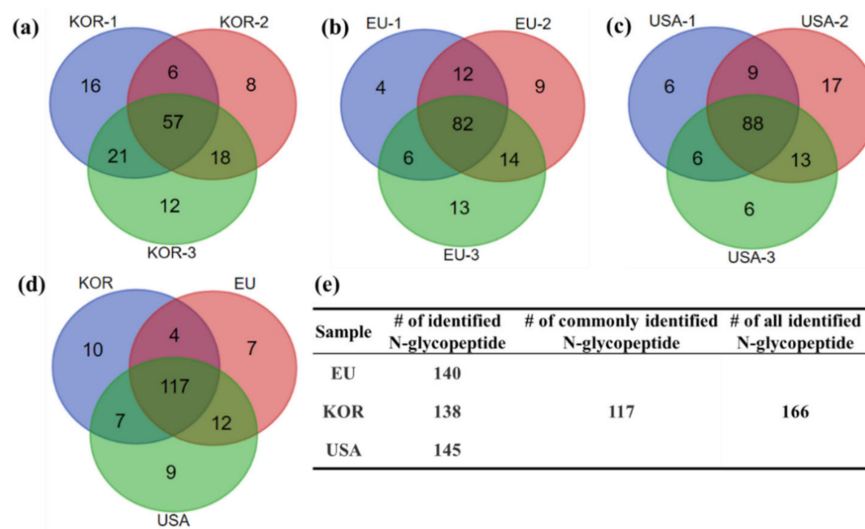
A glycopeptide-based mapping method using liquid chromatography-electrospray ionization tandem mass spectrometry (LC-ESI MS/MS) can be a powerful alternative for the glycosite-specific characterization of complex glycoproteins containing multiple N-glycosylation sites [19–22]. In this study, we used LC-ESI MS/MS for glycopeptide-based mapping of the aflibercept fusion protein. The glycosylation states between the fusion protein lots obtained in different regions were evaluated by site-specifically comparing the glycan microheterogeneity in each N-linked glycosylation site.

## 2. Results and Discussion

### 2.1. Identification of Standard Samples (Korea, Europe and USA)

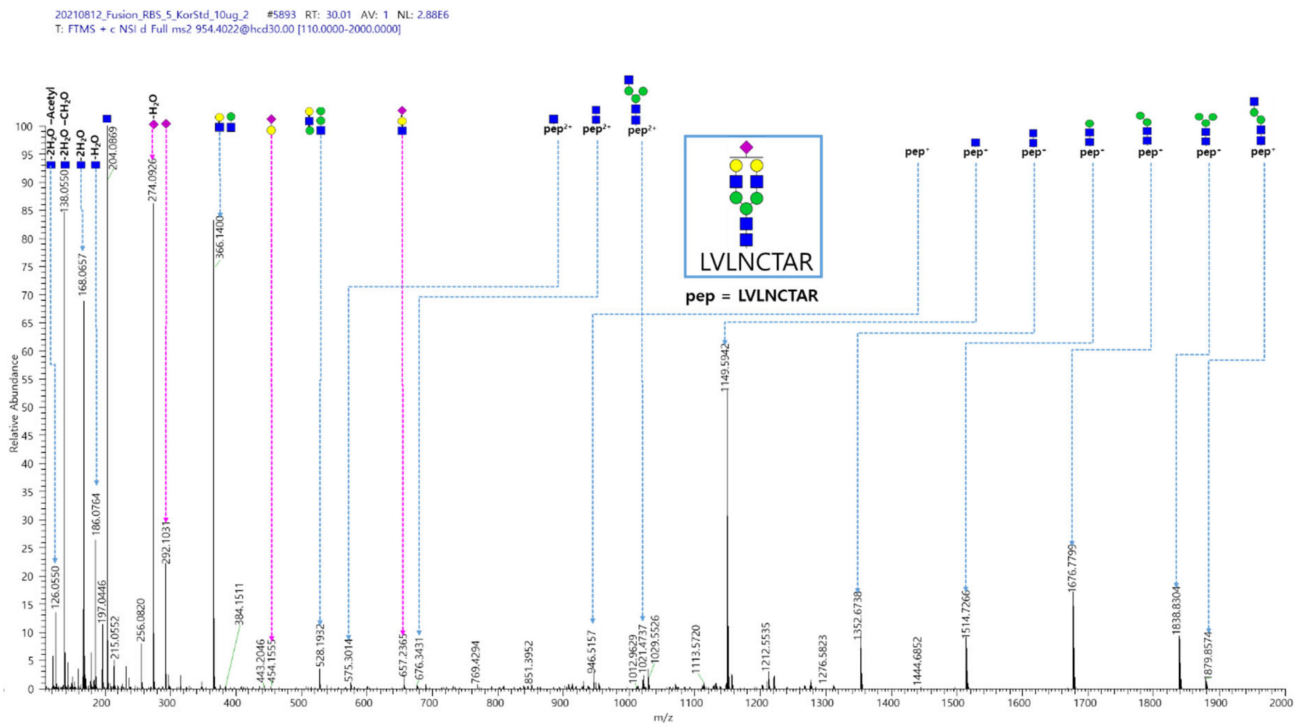
Aflibercept fusion protein (VEGFR-IgG) contains five N-glycosylation sites (two from VEGFR-1, two from VEGFR-2, and one from IgG, as shown in Supplementary Table S1).

VEGFR-IgG fusion protein samples from Korea, Europe and USA (STD-KOR, STD-EU and STD-USA) were tryptic-digested and analyzed by LC-MS/MS with three replicates for the identification and quantification of site-specific glycopeptides. The experimental methods in this work were referenced in our previous study, as the reproducibility of replicate sample preparation and LC-MS/MS analyses for glycoproteins were validated [21]. STD-KOR, STD-EU and STD-USA samples were identified 140, 138 and 145 N-glycopeptides, respectively (Supplementary Table S2) from three replicate LC-MS/MS analyses. The number of total identified N-glycopeptides are 166 (Supplementary Table S3 and over 80% (117) of them were commonly identified (Figure 1).

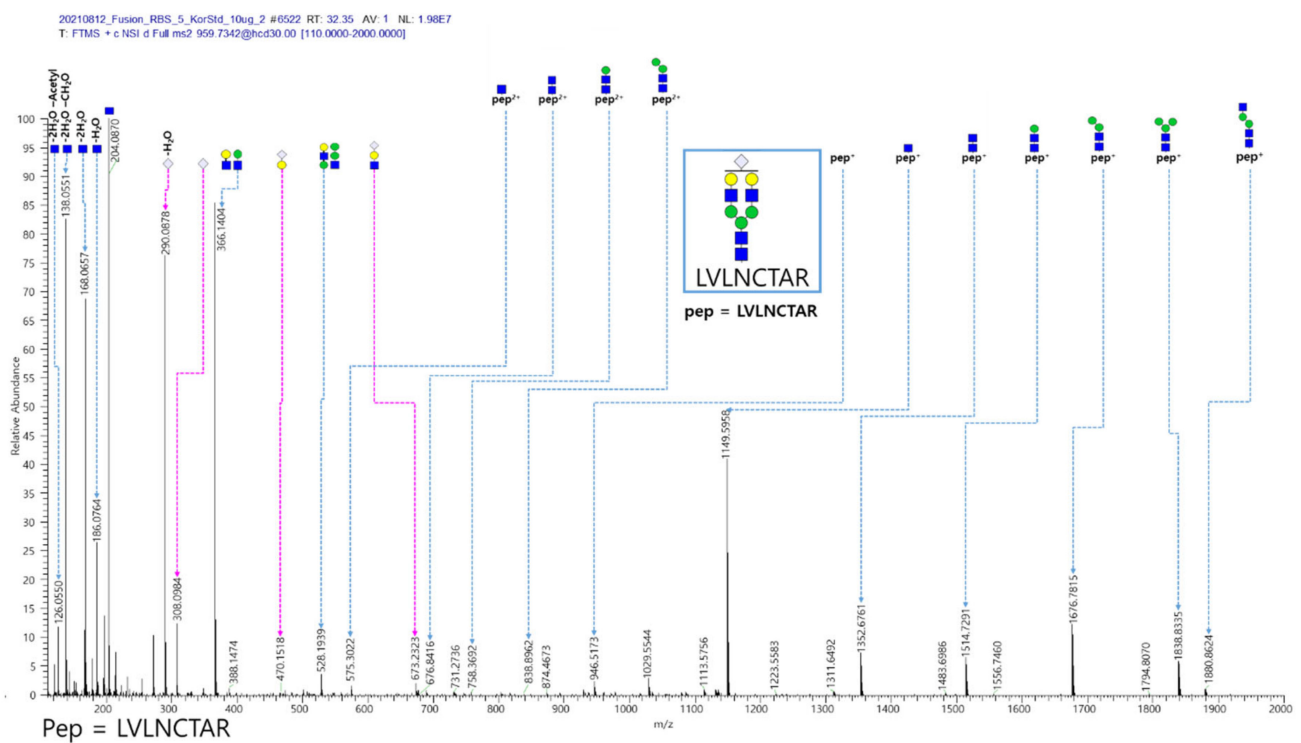


**Figure 1.** N-glycopeptides identified from I-GPA in (a) Korea, (b) Europe, (c) USA and (d) all three standard samples. (e) The summary table of the number of identified N-glycopeptides in each sample.

Animal cells, such as Chinese hamster ovary cells (CHO cells), are generally used for producing mAb [23], thus enabling galactose- $\alpha$ -1,3-galactose ( $\alpha$ -gal) and N-glycolylneuraminic acid (NeuGc), two non-human glycan structures, to be expressed on mAb [24] at very trace levels. They were negligible because of the anaphylactic responses mediated by anti- $\alpha$ -gal IgE [25] and immune reactions mediated by anti-NeuGc antibodies [26] but were difficult to detect due to their very small level. Here, several N-glycopeptides, including NeuGc glycans identified by I-GPA, were expressed in the VEGFR-1 and -2 domains in three standard samples from Korea, Europe and USA, but not in Fc-IgG domain (Supplementary Table S4). The HCD, CID and EThcD spectra from LVLNCTAR\_5\_4\_0\_1\_0 (NeuAc) and LVLNCTAR\_5\_4\_0\_0\_1 (NeuGc) identified by I-GPA search are shown in Figure 2 with oxonium ions from HCD, glycan-cleaved glycopeptide fragment ions (B/Y) from HCD and CID, and peptide fragment ions (c/z) from EThcD. The (a)-1, (b)-1, and (c)-1 spectra in Figure 2 contain NeuAc specific oxonium ions ( $m/z$  274.0926 (NeuAc-H<sub>2</sub>O), 292.1031 (NeuAc), 454.1559 (Hex1NeuAc1), 657.2365 (Hex1HexNAc1NeuAc1)), glycan-cleaved glycopeptide fragment ions ( $m/z$  1167.5120 (Hex3HexNAc3NeuAc1, 2+) and 1248.5358 (Hex4HexNAc3NeuAc1, 2+)) and peptide fragment ions ( $m/z$  1259.9773 (z52+), 1316.5079 (z62+) and 1366.0676 (z72+)) for LVLNCTAR\_5\_4\_0\_1\_0 (NeuAc). The (a)-2, (b)-2, and (c)-2 spectra in Figure 2 contains the NeuGc specific oxonium ions ( $m/z$  290.0878 (NeuGc-H<sub>2</sub>O), 308.0984 (NeuGc), 470.1518 (Hex1NeuGc1), 673.2323 (Hex1HexNAc1NeuGc1)), glycan-cleaved glycopeptide fragment ions ( $m/z$  1175.0062 (Hex3HexNAc3NeuGc1, 2+) and 1256.0327 (Hex4HexNAc3NeuGc1, 2+)) and peptide fragment ions ( $m/z$  1268.1027 (z52+), 1324.5109 (z62+) and 1374.0419 (z72+)) that were identified for LVLNCTAR\_5\_4\_0\_0\_1 (NeuGc). The mass difference between NeuAc and NeuGc specific ions is the mass of oxygen.



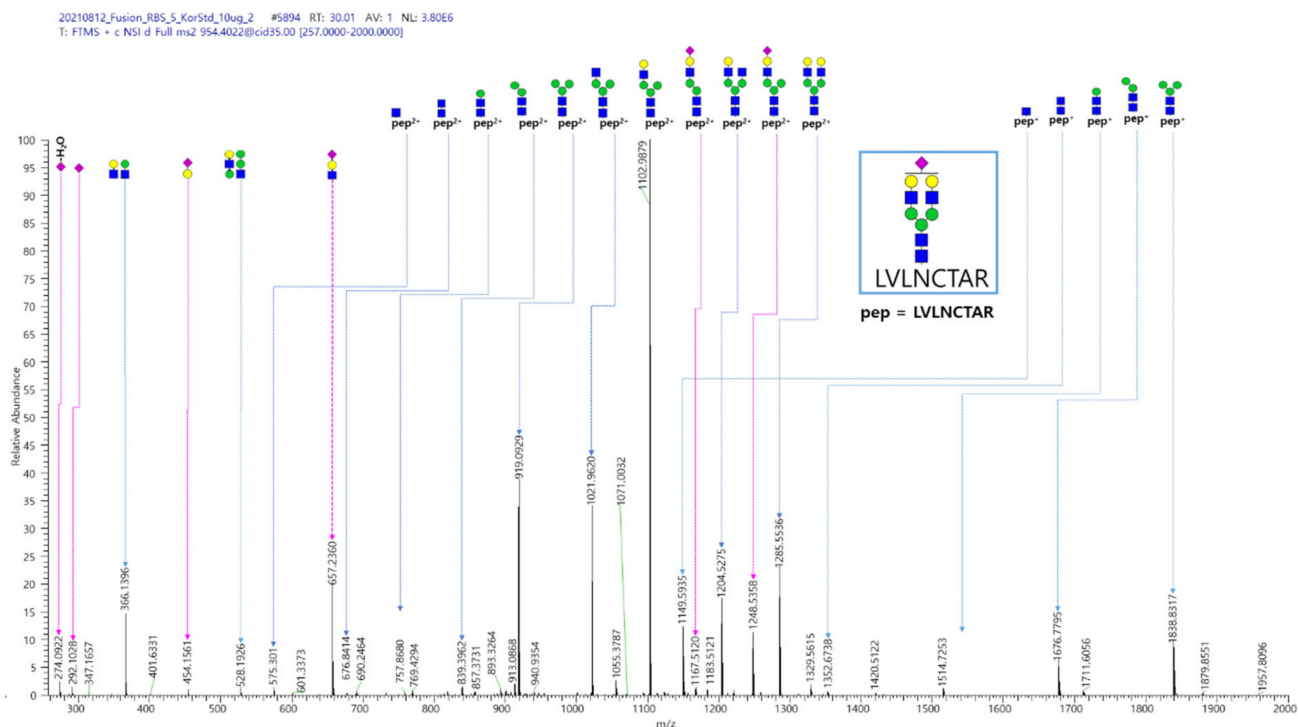
1 HCD spectrum of LVLNCTAR\_5\_4\_0\_1\_0 (NeuAc) N-glycopeptide



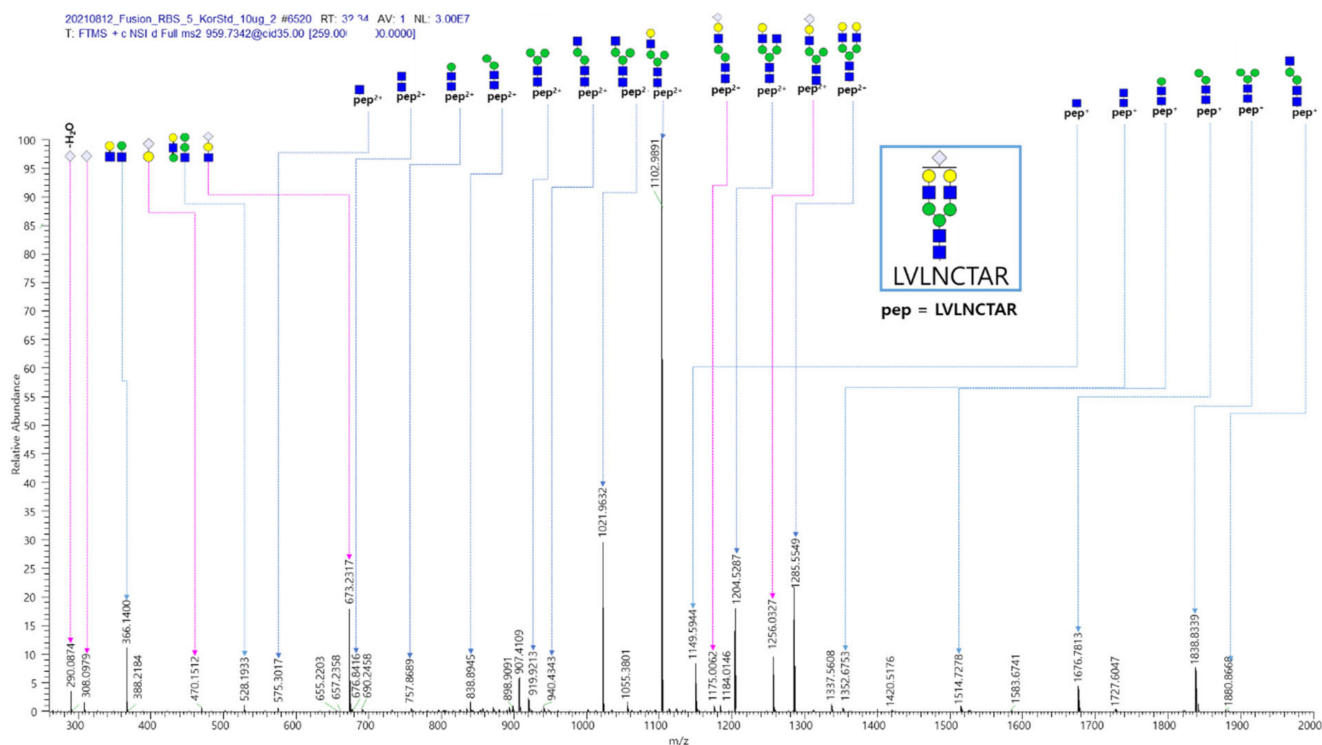
(a) 2 HCD spectrum of LVLNCTAR\_5\_4\_0\_0\_1 (NeuGc) N-glycopeptide

Figure 2. Cont.



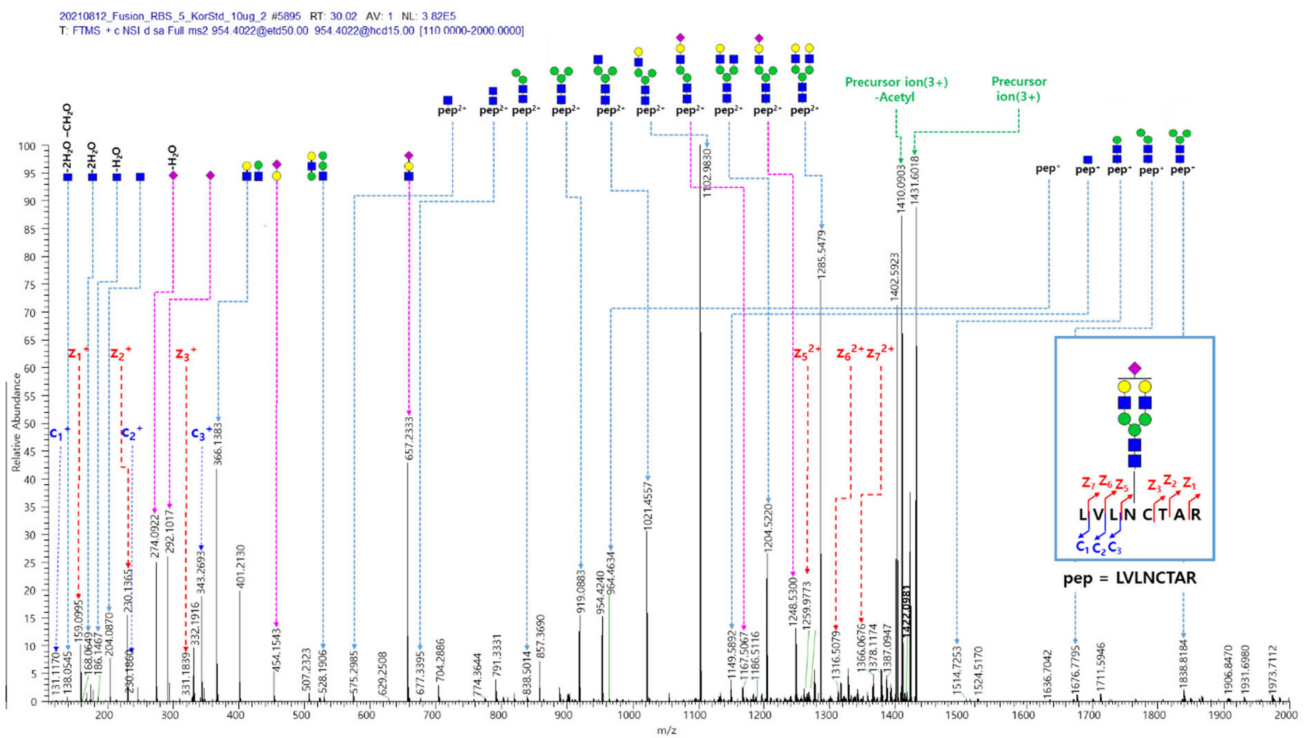


1 CID spectrum of LVLNCTAR\_5\_4\_0\_1\_0 (NeuAc) N-glycopeptide

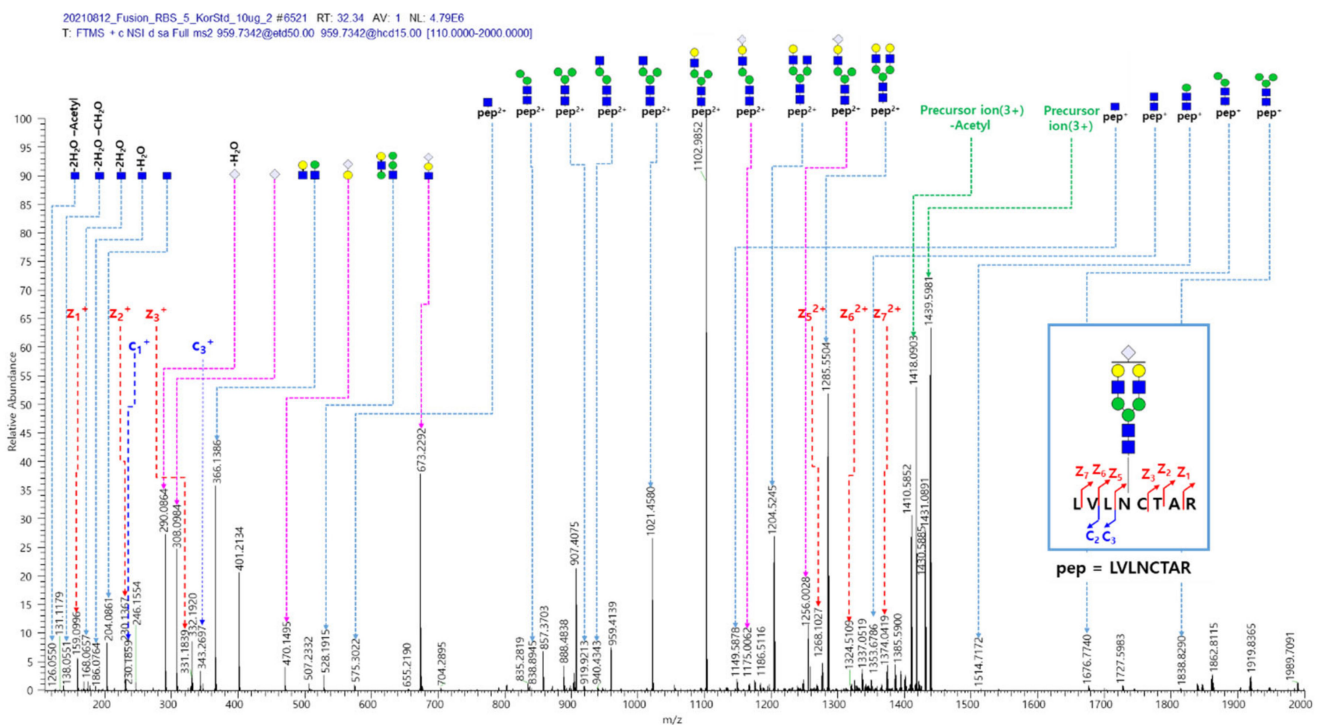


(b) 2 CID spectrum of LVLNCTAR\_5\_4\_0\_0\_1 (NeuGc) N-glycopeptide

Figure 2. Cont.



1 EthcD spectrum of LVLNCTAR\_5\_4\_0\_1\_0 (NeuAc) N-glycopeptide

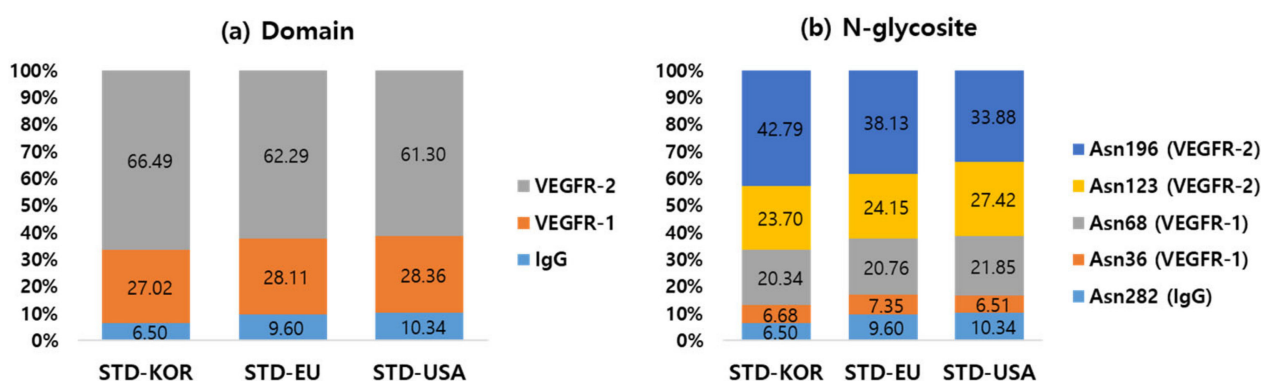


(c) 2 EthcD spectrum of LVLNCTAR\_5\_4\_0\_0\_1 (NeuGc) N-glycopeptide

**Figure 2.** (a) HCD, (b) CID, and (c) EthcD spectra for the LVLNCTAR\_5\_4\_0\_1\_0 (NeuAc) and LVLNCTAR\_5\_4\_0\_0\_1 (NeuGc) N-glycopeptides identified from KOR-standard sample by I-GPA search with specific ions (pink dot lines) for NeuAc and NeuGc.

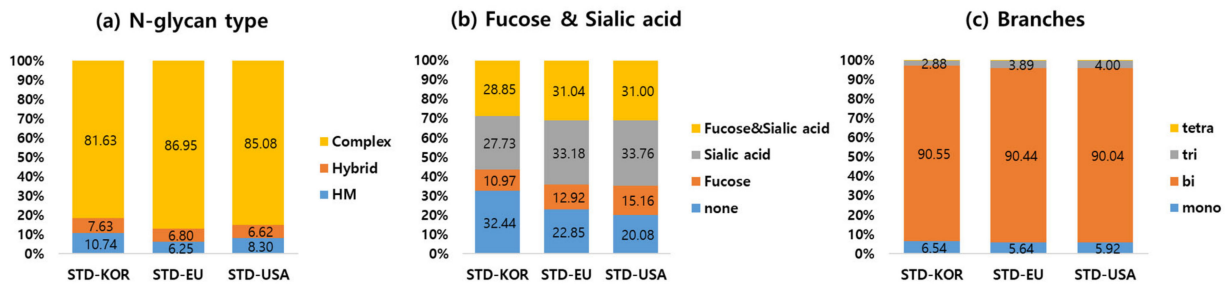
## 2.2. Quantification of Site-Specific N-Glycopeptides from Standard Samples (Korea, Europe and USA)

In total, 166 N-glycopeptides were identified from three LC-MS/MS replications of all three samples using I-GPA (Supplementary Table S3), performed with a q-GPA (quantitated-GPA) [27], for label-free quantitative analysis within the mass tolerance (10 ppm) and the window of retention time (5 min). The quantitative abundance of N-glycopeptides identified at least once in three replicates were extracted from the three replicates data of each sample using q-GPA (Supplementary Figure S1). The intensity of those identified in two or more of the three replicates were averaged and then quantitatively analyzed among samples. Finally, 148, 149, and 148 N-glycopeptides from KOR, EU, and USA standards samples, respectively, were quantified from 155 N-glycopeptides (Supplementary Table S5) and the quantitative results were compared according to the domains (Figure 3a) and the site specific N-glycosites (Figure 3b) in each sample. The abundance of N-glycosylation was approximately 61~66% for VEGFR-2, 27~28% for VEGFR-1, and 7~10% for IgG domain. The site-specific abundance of N-glycosite was arranged Asn196 of VEGFR-2 (34~43%), Asn123 of VEGFR-2 (24~27%), Asn68 of VEGFR-1 (20~22%), Asn36 of VEGFR-1 (7%), and Asn282 of IgG (7~10%).



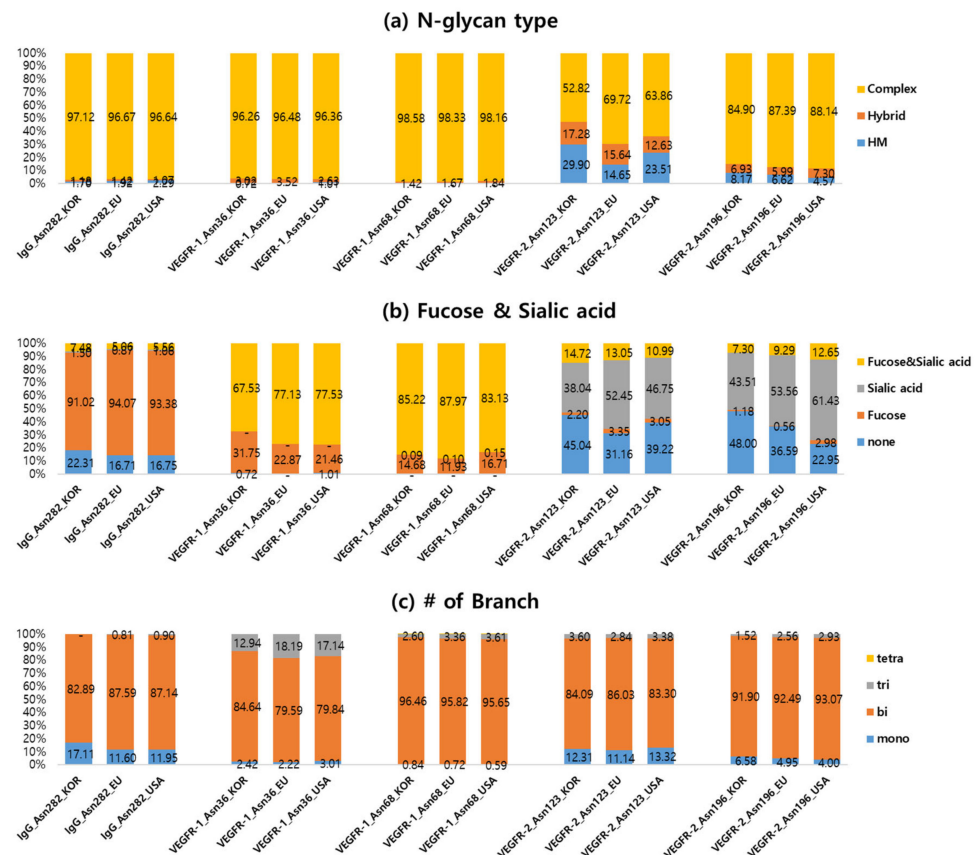
**Figure 3.** The quantitative distributions were compared according to (a) domains and (b) site specific N-glycosites among KOR, EU, and USA standards samples.

Figure 4 shows the distribution according to (a) N-glycan types, (b) fucose or sialic acid glycans strongly related to the protein stability, and (c) the number of branches of attached glycans in KOR, EU, and USA standard samples. There are three main types of N-glycans: high-mannose (HM), complex, and hybrid glycans. High-mannose is consisted of two HexNAc glycans with many mannoses. Complex contains almost any number of the other glycans (Hexose, Fucose, Sialic acid, HexNAc) with two HexNAc glycans. Hybrid glycans have high mannose on one side of the branch and complex glycans on the other side. The most abundant N-glycan was the complex type, comprising over 80% of those identified in all standard samples (Figure 4a). The percentage of complex type in STD-KOR is less than that of other two samples. The intensity of the complex glycans with fucose & sialic acid is almost similar to that of the complex glycans with only sialic acid. The percentages of fucose and sialic acid, sialic acid, and fucose complex glycans in STD-KOR are less than in other two samples. The percentage of complex type without fucose and sialic acid in STD-KOR was greater than in STD-EU and STD-USA. Complex N-glycan types can also be classified according to the number of branches, such as mono-, bi-, tri-, and tetra-antennary. All standards samples have about 90% bi-antennary complex glycans.



**Figure 4.** The quantitative distributions were compared according to (a) N-glycan type, (b) the inclusion of fucose or sialic acid, and (c) the number of branches from complex type glycans in KOR, EU, and USA standards samples.

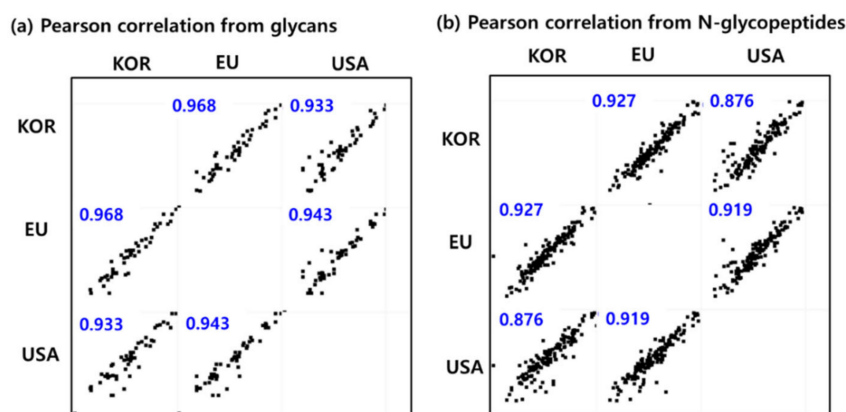
Figure 5 shows the distribution of site-specific N-glycosylation according to the type of N-glycan, the complex glycans with fucose and sialic acid, and the number of branches. IgG and VEGFR-1 domains contain over 95% of the complex glycan types, but the VEGFR-2 domain was approximately 53~70% and 85~88% complex type glycans for Asn123 and Asn196, respectively. Asn123 has the most high-mannose type glycans (15~30%) among the five N-glycosites (Figure 5a). IgG dominantly has bi-antennary fucosylated glycans, while VEGFR-1 highly contains bi-antennary fucosylated and sialylated glycans. VEGFR-2 consisted of 38~61% sialylated and 23~45% non-fucosylated/sialylated glycans that were bi-antennary. IgG and VEGFR-2 have about 4~17% mono-branched glycans, while VEGFR-1 has approximately 3~18% tri- or tetra-branched glycans.



**Figure 5.** The quantitative distribution for each N-glycosite was compared according to (a) N-glycan type, (b) the inclusion of fucose or sialic acid, and (c) the number of branches from complex type glycans in KOR, EU, and USA standards samples.



Finally, the similarity among the three standard samples was represented according to quantitative glycans and glycopeptides with their Pearson correlation values (Figure 6). Figure 6a,b represent the Pearson correlations from 54 glycans (Supplementary Table S6) and 155 N-glycopeptides (Supplementary Table S3), respectively, between the three samples. The similarity between KOR and EU samples, between KOR and USA samples, and between EU and USA samples were 0.968, 0.933, and 0.943, respectively, based on the glycan quantitative analysis, and 0.927, 0.876, and 0.919, respectively, according to the site-specific N-glycopeptide quantitative results. Therefore, the quantification results from the site-specific N-glycopeptides provide us with detailed information regarding the similarity among the fusion protein samples.



**Figure 6.** The Pearson correlations from (a) 54 glycans and (b) 155 N-glycopeptides between three samples.

### 3. Materials and Methods

#### 3.1. Materials

Amicon<sup>®</sup> Ultra centrifugal filters was purchased from Merck Millipore (Burlington, MA, USA). A C18 trap column from Harvard Apparatus (Holliston, MA, USA) was used. Trypsin from Promega (Madison, WI, USA) was used for protein digestion. Phosphate buffered saline (PBS), 1,4-dithiothreitol (DTT) and iodoacetamide (IAA) from Sigma Aldrich (St. Louis, MO, USA) were obtained for reduction and alkylation. Trifluoroacetic acid (TFA), ammonium bicarbonate (ABC), and formic acid (FA) were also purchased from Sigma Aldrich (St. Louis, MO, USA). Water and acetonitrile (ACN) (MS-grade) were purchased from Merck Millipore (Germany)

#### 3.2. Aflibercept Fusion Protein Samples

A total of three different lots of Eylea samples were purchased from the US, Korea, and Europe. Each fusion protein sample was desalted and concentrated using Amicon<sup>®</sup> Ultra centrifugal filters with a PBS buffer.

#### 3.3. Digestion of Fusion Protein

Each protein sample (one hundred  $\mu$ g) with 50 mM ABC was reduced by 10 mM DTT at 60 °C with 850 rpm for 1 hr, alkylated with 25 mM IAA in a darkroom for 1 hr, and digested with trypsin (0.125  $\mu$ g/  $\mu$ L, protein:enzyme = 20:1 (wt/wt)) at 37 °C for 16 hr. The digests were dried by vacuum centrifugation and stored at  $-20$  °C until LC-MS/MS analysis was performed.

#### 3.4. LC-MS/MS Analysis

The digested sample was reconstituted in DW with 0.1% FA and analyzed by an LC-MS/MS system equipped with an Easy nLC 1200 (Thermo Fisher Scientific, USA) and Orbitrap Fusion Lumos mass spectrometer (Thermo Fisher Scientific) consisting of a nano-electrospray ion (ESI) source (EASY-Spray Sources, Thermo Fisher Scientific). For

trapping and separating peptides, a 75  $\mu\text{m} \times 2\text{ cm}$   $\text{C}_{18}$  Pre-column (nanoViper, Acclaim PepMap<sup>TM</sup>100, Thermo Fisher Scientific) and an analytical  $\text{C}_{18}$  column (75  $\mu\text{m} \times 50\text{ cm}$  PepMap<sup>TM</sup> RSLC, Thermo Fisher Scientific) were used, respectively. Flow rate was set at 250 nL/min. Water with 0.1% FA and with 80% acetonitrile, were the mobile phases A and B, respectively. For the LC gradient, 1.6 % B ramped to 4.8% B for 1 min, increased to 8% B (over 12 min), to 28% B (over 60 min), and to 80% B (over 1 min). 80% B was maintained over 5 min, and decreased to 1.6% B for another 3 min. For re-equilibration of the analytical column, 1.6% B was maintained for 8 min before the next run. Approximately 1900 V was the set voltage for producing an electrospray. During the chromatographic separation, the Orbitrap Fusion Lumos was set in data-dependent mode, MS1 and MS2 automatically switched with 3 s of cycle time. The MS data were set using the following parameters: full scan MS1 range was set from 400 to 2500 m/z and acquired by Orbitrap for a maximum ion injection time of 100 ms at a resolution of 120,000 and an automatic gain control (AGC) with a target value of  $1.0 \times 10^6$ ; for MS2 spectra, high energy collision dissociation (HCD) and collision induced dissociation (CID) were performed in the Orbitrap mass analyzer at a resolution of 30,000, each with 35% and 30% normalized collision energy and an AGC target value of  $1.0 \times 10^5$  with a maximum ion injection time of 150 ms. Previously fragmented ions were excluded for 30 s with 10 ppm mass tolerance. The internal calibration was conducted with the mass peak set at 445.12003 m/z, which released from polysiloxane from the silica capillary of the NanoSprayer [28].

### 3.5. N-Glycopeptide Identification

An Integrated GlycoProteome Analyzer (I-GPA) was used to automatically identify N-glycosite-specific glycopeptides from VEGFR-IgG fusion proteins [27]. In summary, N-glycopeptide tandem spectra were first selected with 15 specific oxonium ions of N-glycans. N-glycopeptide candidates were selected for the similarity between their experimental and theoretical MS isotope patterns of N-glycopeptides in the GPA-DB. This GPA-DB combined possible tryptic peptides having N-glycosites of VEGFR-IgG proteins with 351 N-glycans (331 retrosynthetic glycans [29] and 20 polylactosamine series glycans, such as penta- and hexa- [30]). Finally, N-glycopeptides were identified by their Y-score, which represented the degree of matching of the experimental and theoretical fragment ions. From HCD and CID ms/ms spectra, B/Y ions produced from the glycosidic bond cleavage of N-glycans from N-glycosylation sites (oxonium ions and glycan cleaved glycopeptide fragment ions) and b/y ions from peptide bond cleavage (peptide backbone fragment ions), were considered for Y-scoring. The I-GPA search parameters fixed the carbamidomethyl cysteine modification and the mammalian N-glycan for N-sites with one missed tryptic cleavage for peptides. The mass tolerances for precursor and for HCD fragment ions (CID and ETD) were set at 0.02 and 0.02 (each 0.05) Da, respectively. N-glycopeptides within a 0.01 Y-score of FDR were identified from a single targeted glycoprotein.

The nomenclature of the N-glycopeptides was based on the peptide sequence and numbers of hexoses (Hex), GlcNAc, Fuc, N-acetylneuraminic acid (NeuAc), and N-glycolylneuraminic acid (NeuGc) (#Hex\_#HexNac\_#Fuc\_#NeuAc\_#NeuGc). The nomenclature of the N-glycans was based on a retrosynthetic glycans database [20]. For example, EEQYNSTYR\_4\_4\_1\_0\_0 N-glycopeptide was composed of EEQYNSTYR (the peptide sequence), 4 (the number of Hex), 4 (the number of N-GlcNAc), 1 (the number of Fuc), 0 (the number of NeuAc), and 0 (the number of NeuGc). The italicized and underlined N (N) represents an N-glycosylation site.

### 3.6. Glycopeptide Quantification

For label free quantitative analysis of N-glycopeptides, the q-GPA (quantitated-GPA) calculated by the I-GPA software [27] was used as the sum of the three isotopes of the three most intense MS1s. The quantitative abundances of all N-glycopeptides identified from three LC-MS/MS replicates of STD-KOR, STD-EU, and STD-USA, were extracted by q-GPA based on the mass tolerance (10 ppm) and the window of retention time (5 min) of unidentified precursor ions. The intensity from at least two of the three replicates were

averaged and then quantitatively analyzed among samples. Finally, uniquely quantified N-glycopeptides were manually inspected, after the retention time had elapsed, for their isotopic mass distribution pattern.

#### 4. Conclusions

Through the high through-put nano-LC MSMS analysis of the tryptic-digested samples of aflibercept therapeutics obtained from three different local sites, the protein glycosylation states between the aflibercept therapeutics were compared based on the site-specific N-linked glycan microheterogeneity mapping. The glycopeptide-based mapping results of the five N-glycosylation sites showed high similarity between three aflibercept samples when the mapping data was compared statistically at glycan level. However, in site-specific comparison at glycopeptide level based on the identified 155 N-glycopeptides, notable variances were observed between the aflibercept samples. These decreased correlations between samples mainly originated from the variations in the glycan type and the sialylation level on the two N-glycosylation sites in the VEGFR-2 domain, whereas the other N-glycosylation sites (two sites in the VEGFR-1 region and one site in the human IgG Fc region) showed high similarity between the three aflibercept samples in their glycan microheterogeneity.

In addition, it should be noted that several N-glycopeptides, including NeuGc glycan, were identified from the glycopeptide-based mapping of the three aflibercept samples. The glycopeptides possessing NeuGc glycans were mainly identified on VEGFR1 and VEGFR2 domains, and not on the IgG Fc domain. Since it is supposed that NeuGc glycans mediate adverse immune reactions in antibody therapeutics, the presence of these is not negligible. These results imply that much attention is required to validate the similarity and guarantee the quality of fusion protein therapeutics, and this study shows that the site-specific glycopeptide-based mapping approach is a powerful tool for evaluating fusion protein therapeutics and in developing its biosimilars.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms231911807/s1>. Figure S1. Quantitative analysis of 166 N-glycopeptides identified from KOR, EU and USA standard samples, by using q-GPA method. Table S1. The information of VEGFR-IgG fusion protein. Table S2. Site specific N-glycopeptides identified among three samples. Table S3. Site-specifically identified all N-glycopeptides from three samples. Table S4. NeuGc-N-glycopeptide identified by each sample. Table S5. The 155 site specific N-glycopeptide quantitative results among three samples. Table S6. The 54 glycans quantitative results among three samples.

**Author Contributions:** Conceptualization, Y.H.A.; methodology, J.Y.L.; validation, J.-W.C. and J.Y.L.; formal analysis, J.-W.C. and S.H.; data curation, S.H. and J.Y.L.; writing—original draft preparation, Y.H.A. and J.Y.L.; writing—review and editing, S.H.H. and Y.H.A.; supervision, Y.H.A.; project administration, S.H.H. and Y.H.A.; funding acquisition, Y.H.A. and J.Y.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2021R1F1A1061840); and “Regional Innovation Strategy (RIS)” through the NRF funded by the Ministry of Education (MOE) (2021RIS-001); and Bio-convergence Technology Education Program through the Korea Institute for Advancement Technology (KIAT) funded by the Ministry of Trade, Industry and Energy (No. P0017805).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** All data supporting the findings of this study are available within the paper published online.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Schjoldager, K.T.; Narimatsu, Y.; Joshi, H.J.; Clausen, H. Global view of human protein glycosylation pathways and functions. *Nat. Rev. Mol. Cell Biol.* **2020**, *21*, 729–749. [[CrossRef](#)] [[PubMed](#)]
- Stowell, S.R.; Ju, T.; Cummings, R.D. Protein glycosylation in cancer. *Annu. Rev. Pathol.* **2015**, *10*, 473–510. [[CrossRef](#)] [[PubMed](#)]
- Van Landuyt, L.; Lonigro, C.; Meuris, L.; Callewaert, N. Customized protein glycosylation to improve biopharmaceutical function and targeting. *Curr. Opin. Biotechnol.* **2019**, *60*, 17–28. [[CrossRef](#)] [[PubMed](#)]
- Berger, M.; Kaup, M.; Blanchard, V. Protein glycosylation and its impact on biotechnology. *Adv. Biochem. Eng. Biotechnol.* **2012**, *127*, 165–185. [[CrossRef](#)]
- Hajba, L.; Szekrényes, Á.; Borza, B.; Guttman, A. On the glycosylation aspects of biosimilarity. *Drug Discov. Today* **2018**, *23*, 616–625. [[CrossRef](#)]
- Ackermann, M.; Jäger, V.; Marx, U. Influence of cell- and media-derived factors on the integrity of a human monoclonal antibody after secretion into serum-free cell culture supernatants. *Biotechnol. Bioeng.* **1995**, *45*, 97–106. [[CrossRef](#)]
- Hennicke, J.; Reinhart, D.; Altmann, F.; Kunert, R. Impact of temperature and pH on recombinant human IgM quality attributes and productivity. *New Biotechnol.* **2019**, *50*, 20–26. [[CrossRef](#)]
- Parhiz, H.; Ketcham, S.A.; Zou, G.; Ghosh, B.; Fratz-Berilla, E.J.; Ashraf, M.; Ju, T.; Madhavarao, C.N. Differential effects of bioreactor process variables and purification on the human recombinant lysosomal enzyme  $\beta$ -glucuronidase produced from Chinese hamster ovary cells. *Appl. Microbiol. Biotechnol.* **2019**, *103*, 6081–6095. [[CrossRef](#)]
- Sissolak, B.; Lingg, N.; Sommeregger, W.; Striedner, G.; Vorauer-Uhl, K. Impact of mammalian cell culture conditions on monoclonal antibody charge heterogeneity: An accessory monitoring tool for process development. *J. Ind. Microbiol. Biotechnol.* **2019**, *46*, 1167–1178. [[CrossRef](#)]
- Edwards, E.; Livanos, M.; Krueger, A.; Dell, A.; Haslam, S.M.; Mark Smales, C.; Bracewell, D.G. Strategies to control therapeutic antibody glycosylation during bioprocessing: Synthesis and separation. *Biotechnol. Bioeng.* **2022**, *119*, 1343–1358. [[CrossRef](#)]
- Mimura, Y.; Saldova, R.; Mimura-Kimura, Y.; Rudd, P.M.; Jefferis, R. *Importance and Monitoring of Therapeutic Immunoglobulin G Glycosylation*; Experientia Supplementum; Springer: Cham, Switzerland, 2021; Volume 112, pp. 481–517. [[CrossRef](#)]
- Dolashka, P. Tandem mass spectrometry and glycoproteins. In *Tandem Mass Spectrometry—Applications and Principles*; Prasain, J., Ed.; Intechopen: London, UK, 2012; Chapter 5; pp. 105–126.
- An, H.J.; Froehlich, J.W.; Lebrilla, C.B. Determination of glycosylation sites and site-specific heterogeneity in glycoproteins. *Curr. Opin. Chem. Biol.* **2009**, *13*, 421–426. [[CrossRef](#)] [[PubMed](#)]
- Liu, L. Antibody Glycosylation and Its Impact on the Pharmacokinetics and Pharmacodynamics of Monoclonal Antibodies and Fc-Fusion Proteins. *J. Pharm. Sci.* **2015**, *104*, 1866–1884. [[CrossRef](#)]
- Higel, F.; Seidl, A.; Sörgel, F.; Friess, W. N-glycosylation heterogeneity and the influence on structure, function and pharmacokinetics of monoclonal antibodies and Fc fusion proteins. *Eur. J. Pharm. Biopharm.* **2016**, *100*, 94–100. [[CrossRef](#)] [[PubMed](#)]
- Ruhaak, L.R.; Zauner, G.; Huhn, C.; Bruggink, C.; Deelder, A.M.; Wuhrer, M. Glycan labeling strategies and their use in identification and quantification. *Anal. Bioanal. Chem.* **2010**, *397*, 3457–3481. [[CrossRef](#)] [[PubMed](#)]
- Bongers, J.; Devincendis, J.; Fu, J.; Huang, P.; Kirkley, D.H.; Leister, K.; Liu, P.; Ludwig, R.; Rumney, K.; Tao, L.; et al. Characterization of glycosylation sites for a recombinant IgG1 monoclonal antibody and a CTLA4-Ig fusion protein by liquid chromatography–mass spectrometry peptide mapping. *J. Chromatogr. A* **2011**, *1218*, 8140–8149. [[CrossRef](#)] [[PubMed](#)]
- Ciombor, K.K.; Berlin, J.; Chan, E. Aflibercept. *Clin. Cancer Res.* **2013**, *19*, 1920–1925. [[CrossRef](#)]
- Hahm, Y.H.; Lee, J.Y.; Ahn, Y.H. Investigation of Site-Specific Differences in Glycan Microheterogeneity by N-Glycopeptide Mapping of VEGFR-IgG Fusion Protein. *Molecules* **2019**, *24*, 3924. [[CrossRef](#)]
- Zhu, L.; Guo, Q.; Guo, H.; Liu, T.; Zheng, Y.; Gu, P.; Chen, X.; Wang, H.; Hou, S.; Guo, Y. Versatile characterization of glycosylation modification in CTLA4-Ig fusion proteins by liquid chromatography-mass spectrometry. *mAbs* **2014**, *6*, 1474–1485. [[CrossRef](#)]
- Lee, J.Y.; Lee, H.K.; Park, G.W.; Hwang, H.; Jeong, H.K.; Yun, K.N.; Ji, E.S.; Kim, K.H.; Kim, J.S.; Kim, J.W.; et al. Characterization of Site-Specific N-Glycopeptide Isoforms of  $\alpha$ -1-Acid Glycoprotein from an Interlaboratory Study Using LC-MS/MS. *J. Proteome Res.* **2016**, *15*, 4146–4164. [[CrossRef](#)]
- Reusch, D.; Habberger, M.; Falck, D.; Peter, B.; Maier, B.; Gassner, J.; Hook, M.; Wagner, K.; Bonnington, L.; Bulau, P.; et al. Comparison of methods for the analysis of therapeutic immunoglobulin G Fc-glycosylation profiles—Part 2: Mass spectrometric methods. *mAbs* **2015**, *7*, 732–742. [[CrossRef](#)]
- Read, E.K.; Park, J.T.; Brorson, K.A. Industry and regulatory experience of the glycosylation of monoclonal antibodies. *Biotechnol. Appl. Biochem.* **2011**, *58*, 213–219. [[CrossRef](#)] [[PubMed](#)]
- Ghaderi, D.; Zhang, M.; Hurtado-Ziola, N.; Varki, A. Production platforms for biotherapeutic glycoproteins. Occurrence, impact, and challenges of non-human sialylation. *Biotechnol. Genet. Eng. Rev.* **2012**, *28*, 147–175. [[CrossRef](#)] [[PubMed](#)]
- Chung, C.H.; Mirakhur, B.; Chan, E.; Le, Q.T.; Berlin, J.; Morse, M.; Murphy, B.A.; Satinover, S.M.; Hosen, J.; Mauro, D.; et al. Cetuximab-induced anaphylaxis and IgE specific for galactose- $\alpha$ -1,3-galactose. *N. Engl. J. Med.* **2008**, *358*, 1109–1117. [[CrossRef](#)] [[PubMed](#)]
- Ghaderi, D.; Taylor, R.E.; Padler-Karavani, V.; Diaz, S.; Varki, A. Implications of the presence of N-glycolylneuraminic acid in recombinant therapeutic glycoproteins. *Nat. Biotechnol.* **2010**, *28*, 863–867. [[CrossRef](#)]



27. Park, G.W.; Kim, J.Y.; Hwang, H.; Lee, J.Y.; Ahn, Y.H.; Lee, H.K.; Ji, E.S.; Kim, K.H.; Jeong, H.K.; Yun, K.N.; et al. Integrated GlycoProteome Analyzer (I-GPA) for Automated Identification and Quantitation of Site-Specific N-Glycosylation. *Sci. Rep.* **2016**, *6*, 21175. [[CrossRef](#)]
28. Keller, B.O.; Sui, J.; Young, A.B.; Whittall, R.M. Interferences and contaminants encountered in modern mass spectrometry. *Anal. Chim. Acta* **2008**, *627*, 71–81. [[CrossRef](#)]
29. Kronewitter, S.R.; An, H.J.; de Leoz, M.L.; Lebrilla, C.B.; Miyamoto, S.; Leiserowitz, G.S. The development of retrosynthetic glycan libraries to profile and classify the human serum N-linked glycome. *Proteomics* **2009**, *9*, 2986–2994. [[CrossRef](#)]
30. Ozohanic, O.; Turiak, L.; Puerta, A.; Vekey, K.; Drahos, L. High-performance liquid chromatography coupled to mass spectrometry methodology for analyzing site-specific N-glycosylation patterns. *J. Chromatogr. A* **2012**, *1259*, 200–212. [[CrossRef](#)]