MEDDOCS Open Access Publisher

ISSN: 2637-4927

Annals of Biotechnology

Open Access | Research Article

Two-peak Elution Profile of a Bispecific VHH-IgG Fusion Protein in Ion Exchange Purification Process caused by Atypical N-Glycosylation

*Jing Jiao[1]; Liming Huang[1]; Xiaohui Xu[1]; Yuying Liu[2]; Mingyan Hu[1]; Weifeng Chen[1]; Lei Li[1]; Hui Xu[1]; Xugang He[2]; Nannan Li[2]; Aihua Chen[1]; Chenxi Shen[2]; Lei Wang[2]; Xun Liu[3]; Lianshan Zhang[4]; Chenxiang Tang[1]**

[1]*Shanghai MabGen Biotech Ltd, China.*

[2]*Beijing Tuojie Biotech Co Ltd, China.*

[3]*Shanghai Hengrui Pharmaceuticals Co Ltd, China.*

[4]*Jiangsu Hengrui Pharmaceuticals Co Ltd, China.*

***Corresponding Author(s): Chenxiang Tang**

Shanghai Mabgen Biotech LTD, No.18 Jin Dian Road, Pudong district, Shanghai, China.

Tel: +86-21-58880081-8666, Fax: +86-21-58880181;

Email: chenxiang_tang@163.com

Received: Apr 12, 2022

Accepted: May 07, 2022

Published Online: May, 09, 2022

Journal: Annals of Biotechnology

Publisher: MedDocs Publishers LLC

Online edition: http://meddocsonline.org/

Keywords: Atypical N-glycosylation; Asn-X-Cys motif; VHH domain; Bispecific antibody; Two-peak elution profile; Mass spectrometry.

Abstract

In-depth analysis and thorough understanding of N-glycosylation is critical for therapeutic proteins. For IgG or IgG-like molecules, the conserved N-glycosylation at Asn-297 has critical safety and efficacy implications and therefore is under close scrutiny. Any additional N-glycosylation that increases complexity and heterogeneity of the molecule, should be avoided at the early stage of drug discovery. However, unexpected glycosylation at atypical sites brings unforeseen challenges for antibody drug development process. Here, we observed a two-peak elution profile when developing ion exchange purification process for a novel bispecific VHH-IgG fusion protein. The charge and size properties of the two peaks were characterized by a variety of analytical methods. Two atypical N-glycosylation sites in the VHH domain were identified by means of mass spectrometry and confirmed by site-specific mutagenesis. Further characterization indicated that G0F was the predominant glycan species for both sites with varying occupancy. Interestingly, the addition of neutral glycans changed the charge behaviour of the fusion protein, leading to the unexpected ion exchange profile. These findings suggest that glycosylation may occur at non-canonical sites and may have significant effect on the process development and drug developability.

Introduction

Biologics have emerged as the most fast-growing segment in the pharmaceutical industry. With rapid advancement in the antibody engineering technologies, therapeutic antibodies with different formats, such as Antibody-Drug Conjugates (ADCs), multi-specific antibodies and domain antibody fusions, have been approved by the regulatory agencies worldwide. These new modalities have significant impact on the "manufacturability" of the products [1]. Downstream process for the separation of product from product-related impurities becomes ever more challenging. Abnormal elution profiles of Ion Exchange Chroma-

Cite this article: Jiao J, Huang L, Xu X, Liu Y, Hu M et al. Two-peak Elution Profile of a Bispecific VHH-IgG Fusion Protein in Ion Exchange Purification Process caused by Atypical N-Glycosylation. Ann Biotechnol. 2022; 5(1): 1025.

tography (IEC), such as split peaks, have been seen more frequently in mAb purification especially for mAb-derivatives [2,3]. These IEC profiles are mostly resulted from heterogeneity of the protein molecules. Studies suggest that the potential causes of the heterogeneity include aggregation, fragmentation, glycosylation, charge variation, etc [4,5,6,7].

N-glycosylation profile is largely determined by the sequence of the molecule, the host cell and the manufacturing process. Main glycoforms in IgG-Fc domain are complex-type neutral glycans such as G0F, G1F, G2F. Truncated complex-type glycans (G0F-GN, G1F-GN) and afucosylated glycans (G0, G1, G2) are also found, albeit less often. Additionally, therapeutic proteins with varying glycan profiles may affect safety and efficacy of the product [8,9,10]. Most of the approved therapeutic antibodies carry only one N-glycosylation site Asn-297 in the CH2 domain on each heavy chain. The glycosylation is critical for its therapeutic half-life and for many important biological functions such as Antibody-Dependent Cell Mediated Cytotoxicity (ADCC), Antibody-Dependent Cellular Phagocytosis (ADCP), and Complement-Dependent Cytotoxicity (CDC) [11]. Therefore, a thorough characterization of the protein glycosylation is required for its potential adverse effects on the stability, safety and functionality of the molecule [12].

At present, a number of approved therapeutic antibodies, especially Fc-fusion proteins, have multiple sites of glycosylation [13,14,15,16], often due to lack of information on glycosylation sequon. Normally, N-glycosylation occurs at the Asn-X-Ser/Thr, motif, where X could be any amino acid except proline. Recently, several research groups have identified atypical N-glycosylation motif via optimized sample preparation, higher-sensitivity mass spectrometry and more powerful informatics tools. Mann's group [17] made a comprehensive mapping of an in vivo N-Glycoproteome in mouse tissue and plasma. A total of 6367 N-glycosylation sites on 2352 proteins were identified, among which 112 N-glycosylation sites didn't match consensus motif. Zhang's group [18,19] published a series of papers showing atypical N-glycosylation sequon including Asn-His-Val, Asn-Ser-Cys and Asn-Gly-Val by a novel solid-phase extraction of the N-Linked Glycans and Glycosite-Containing Peptides (NGAG) method. Daisuke Yasuda et al [20] found that human GPR109A has an N-glycosylation motif of Asn-Cys-Cys which plays crucial biologic roles. Six N-glycosylation sites within Asn-X-Cys motif were also identified by Mark S. et al in Alpha-1-Acid Glycoprotein (A1AG) and serotransferrin. Highly N-glycosylated A1AG at Asn-Gln-Cys in porcine and canine sources was found through semi-quantitation based on Multiple-Reaction Monitoring (MRM) analysis [21]. Valliere-Douglass et al. reported N-glycosylation on a non-consensus amino acid sequence in recombinant human IgG1 and IgG2, proving the presence of atypical N-glycosylation sites in recombinant proteins [22,23]. Up until now, few reports regarding atypical N-glycosylation sites in single variable domain (VHH) or VHH-based bispecific antibody have been published.

Here, we report an unexpected double-peak elution profile during the purification process using IEC. We characterized the two peaks using Size Exclusion Chromatography (SEC), Capillary Electrophoresis Sodium Dodecyl Sulfate (CE-SDS), Imaged Capillary Isoelectric Focusing (iCIEF) and mass spectrometry. The results indicate peak-splitting pattern is due to N-glycosylation in the VHH domain. Further study demonstrates that the N-glycosylation modifications are at two atypical sites Asn-Ser-Cys and Asn-Lys-Cys in the VHH domain of a bispecific fusion protein. Our findings, for the first time, identified atypical N-glycosylation motif Asn-X-Cys in VHH domain.

Materials and methods

Sample preparation

Wild type BisAb DNA was synthesized and cloned into PCDNA3.1 vector. The N1Q, N2Q and N1D/N2G mutants were generated by introducing point mutation into wild type PCDNA3.1-BisAb vector at the designated site. We used primeSTAR (premix) (takara, R040) for PCR reactions. The PCR product was digested by DpnI (NEB, R0176L) for 1h, then transformed to DH5a (Biomed, 765666AA). The correct colonies were confirmed by sequencing. Wild type and mutated BisAb were expressed using transiently transfected CHO K1 cells and purified with Protein A column.

Size exclusion chromatography analysis

Analytical size-exclusion chromatography was performed on Agilent 1260 with Xbridge BEH200 SEC column (Waters, 7.8×300 mm, 3.5 µm). The SEC mobile phase included 50 mM sodium phosphate and 200 mM Arginine (pH 6.80). Samples were analyzed at a flow rate of 0.8 mL/minute in an isocratic mode, and chromatographic separation was monitored at 280 nm.

Capillary electrophoresis-sodium dodecyl sulfate analysis

Both reducing and non-reducing CE-SDS analyses were conducted by high-performance capillary electrophoresis system (PA-800 plus Pharmaceutical Analysis System; Beckman). For non-reduced CE-SDS, sample mixed with a 10 kDa internal standard was denatured by 50mM Tris-HCl containing 1% SDS and alkylated by iodoacetamide at 70^{O}C for 10 min. For reduced condition, β-mercaptoethanol was added to reduce the protein disulfide bonds. Samples were injected onto a fused silica capillary applying -5 kV voltage for 20 seconds and separated in the capillary cartridge. All samples run were monitored by 220 nm UV detection.

Imaged Capillary Isoelectric Focusing (iCIEF) analysis

The iCIEF analysis was performed on Maurice (Protein Simple) with a fluorocarbon-coated capillary. Samples were desalted before diluting with sample buffer to a final concentration of 0.2 mg/mL. The sample buffer contained 0.35% methylcellulose, 4% Pharmalyte 3-10, 2 M urea, and 0.05% pI markers (pI 4.05 and 9.50) (Arg and IDA). The focusing was conducted at 1.5 kV for 1 min followed by 3.0 kV for 6 min.

Molecular weight analysis

Waters ACQUITY UPLC (BioResolve RP mAb Polyphenyl column, 2.7 µm 450 Å, 2.1 × 150 mm) coupled to Waters Xevo G2-XS Q-TOF mass spectrometer was used to determine molecular weight. Reduced samples were prepared by incubating with 10 mM Dithiothreitol (DTT, from sigma) at 37°C for 30 min to separate heavy chain and light chain. IdeS digestion was performed according to the protocol of manufacture (50 units IdeS: 1 µg protein) to generate Fd, LC and 1/2Fc fragments. Deglycosylation samples were generated by mixing samples with PNGase F (Rhino Bio) and incubating at 37°C for 30 min. All the prepared samples were diluted to 0.5 mg/ml with ammonium bicarbonate buffer for molecular weight analysis. Two micrograms of each sample were injected and data in the m/z range of 500-4500 Da were acquired. Deconvolution of the MS spectra was realized by Unifi software.

Peptide mapping analysis

Samples were denatured in 8 M guanidine hydrochloride in 50 mM Tris Hydrochloride (Tris-HCl, pH7.0) and 5 mM EDTA buffer for 30min, followed by reduction and alkylation with 5 mM DTT and 10 mM iodoacetamide. Then, samples were buffer-exchanged into 50 mM Tris-HCl at final concentration of 0.5 mg/ml. 50 µg sample was incubated with 2 µg trypsin at 37^{O}C for 30min. Digestion was terminated by adding 2 µL formic acid. The generated peptides were analyzed by Waters ACQUITY UPLC (ACQUITY UPLC Peptide BEH C18 column, 1.7 µm 130 Å, 2.1 mm × 150 mm) coupled to Waters Xevo G2-XS Q-TOF mass spectrometer. Data was processed using Unifi software.

Ion exchange chromatography

Laboratory scale chromatographic experiments were performed on a Cytiva AKTA Avant 150 controlled by Unicorn software version 7.3. The multimodal resin, Capto MMC Impres, was packed into 1.1 cm Inner diameter Vantage® columns or 0.77 cm Inner diameter HiScreenTM prepacked columns, to a bed height of 10±2 cm. The column was equilibrated with 4-5 CV (Column volume) buffer 1 (20 mM sodium phosphate, 10 mM citric acid, adjusted by 1 M Tris, pH6.0). Sample pH was adjusted to 6.0 and then loaded to the column, followed by a 4-5 CV re-equilibration with buffer 1. A 10 CV 0-400 mM NaCl linear gradient elution was performed at pH6.0. The absorbance of protein was monitored at 280 nm.

Results

BisAb exhibits double-peak elution behaviour in IEC

The Bispecific fusion protein in this study (BisAb) consists of a VHH domain fused to the N terminal of a full length Immunoglobuling subclass 4 (IgG4), as shown in Figure 1. The VHH domain and the IgG4 are designed to bind to two different therapeutic targets. IEC was used to purify the BisAb following the initial capture with Protein A column. SEC purity was 97.9% after the affinity purification step (Figure S1, ProA Elution). A split-peak profile was observed during elution step of IEC (Figure 2). Any variants that cause differences in the above properties may lead to the two-peak phenomena, such as aggregation, fragmentation or certain modifications. Fractions of the two peaks in the IEC purification were collected and pooled as BisP1 and BisP2, respectively. Comprehensive analyses were conducted for BisP1 and BisP2 to determine the causes for the peak splits.

BisP1 contains more acidic and larger species

To first determine whether aggregation contributes to the peak split, the size of BisP1 and BisP2 was measured by SEC. As shown in Figure S1, both BisP1 and BisP2 eluted as a monodispersed peak with retention time around 8.2 minutes, corresponding to roughly 175 kD. Although the observed pre- and post- main peak UV absorption suggests the existence of traces of aggregations and fragments, monomer is the predominant species for both BisP1 and BisP2 with over 97% purity. Therefore, it is unlikely that aggregations resulted in the two-peak IEC profile.

We then focused on lower molecular weight species, at which range, CE-SDS provIdes more accurate information. Non-reduced CE-SDS (nrCE-SDS) was performed and results are shown in Figure 3. The main peaks of BisP1 and BisP2 in non-reduced CE-SDS migrate at slightly different time, 28.12 minutes for BisP2 and 28.48 minutes for BisP1. A closer examination reveals that the main peak of BisP1 is accompanied with a left shoulder peak at a level of roughly 6.5%. The shoulder peak elutes at 28.04 minutes, similar to the main peak of BisP2. This suggests that BisP1 contains mixed populations, with the predominant species slightly larger than BisP2.

To achieve better resolution, reduced CE-SDS (rCE-SDS) was conducted on BisP1 and BisP2. As shown in Figure 4, no notable difference was observed for Light Chains (LCs) of BisP1 and BisP2. BisP2 contains a single HC peak at the expected running time of 20.83 minutes. However, in addition to the 20.83-minutes-peak, a second peak showed at 21.38 minutes in BisP1, corresponding to a larger species. The latter peak accounts for about 43.2% of the BisP1 HC. Since each BisAb contains two HCs, up to 86.4% of BisP1 could be the larger species, consistent with nrCE-SDS observation (Figure 3, BisP1). For mAbs or mAb-derivatives, mass increase may be caused by various reasons, including post-translational modifications and sequence variants like mutation, C-terminal Fc-extension or partial leader sequence [24,25].

Understanding whether the identified larger species in BisP1 changes the charge behaviour of BisAb is helpful to narrow down potential causes. iCIEF analysis was hence performed. Although BisP1 and BisP2 share similar Isoelectric point (pI) of 8.5, difference in their charge-variant profiles were significant, shown in Figure 5. Increased acidic (45.9%) and decreased basic (14.0%) populations were detected in BisP1, compared with those in BisP2 (acidic 21.9% and basic 21.6%). The overall acidic shift of BisP1 could explain the initial observation that BisP1 elutes earlier in IEC purification. The results imply that the mass increase may have introduced an acidic change in BisAb.

The mass increase in BisP1 results from glycosylation

To uncover the reason behind the mass increase, the intact Molecular Weight (MW) was determined by high-resolution mass spectrometry, as shown in Figure 6. The theoretical MW of our BisAb with G0F glycan on the conserved Asn-297 site of both IgG4 HCs is 176682 Da. The observed intact MWs for BisP1 (176681 Da) and BisP2 (176679 Da) were in good agreement with the theoretical number. Moreover, mass addition of 1.44K Da was observed in BisP1, but not in BisP2. The rCE-SDS analysis have demonstrated that the impurity with increased mass was higher in BisP1. As expected, the peak of 178124 Da was much higher than the peak of 176681 Da in BisP1 deconvoluted spectra and consecutive 1.44K Da mass shift was observed (Figure 6a). The 1.44k Da mass increase is similar to the glycan species of G0F modification. In addition, characteristic sequential mass difference of 203 Da from two neighbouring peaks in BisP1 indicates possible additions of acetyl hexosamine, further supporting our glycan hypothesis. Our preliminary analysis is in favour of possible glycan addition but not sequence variant.

Reduced molecular weight analysis of BisP1 and BisP2 was then outperformed and the deconvolution results were shown in Figure S2 and Figure S3. The observed MW of HC and LC in both BisP1 and BisP2 is consistent with the theoretical calculation. A cluster of peaks at 66kDa with a mass addition of roughly 1445 Da were detected for BisP1 HC and the mass difference of 203 Da were found again in adjacent peaks. The results were in agreement with both CE-SDS and intact mass analysis, confirming the presence of potential additional glycosylation in the HC of BisP1.

IdeS digestion followed by DTT reduction was applied to further locate the mass change. IdeS digests the HC of IgG4 at the

hinge region, resulting in Fc/2 and Fd+VHH parts (Figure 7a). The deconvolution results of BisP1 and BisP2 were compared in the overlaid spectra. The Fc/2 parts of the two samples share the same observed molecular weight with the theoretical calculation (data not shown). The +1445 Da species was only observed in the Fd+VHH region of BisP1 (Figure 7b), but not in the corresponding region of BisP2 (Figure 7c). The results imply that the impurity was probably composed of Fd+VHH region linked by relatively hydrophilic glycan species. Possible glycan compositions were assigned to the peaks in BisP1 based on the mass, as shown in Figure 7(b). Biantennary structure with core-fucose attachment was the predominant glycan type.

To confirm our hypothesis that additional glycosylation leads to the multiple peaks, deglycosylation reaction with PNGase F followed by DTT reduction was performed. Mass spectrometric analysis reveals that the observed MW values 63645 Da of BisP1 and 63645 Da of BisP2 are highly consisted with theoretical MW of deglycosylated heavy chain (63644 Da), shown in Figure 8. The previous +1445 Da species and the peak clusters separated by 203Da are undetectable after deglycosylation. Our results demonstrate that BisP1 contains product-related impurity that has N-glycan additions to the Fd+VHH region of its HC.

Rare glycosylation at VHH domain leads to two-peak elution IEC profile

Intriguingly, the Fd+VHH part in our molecule does not contain any canonical N-glycosylation motifs (Asn-X-Ser/Thr). To identify the atypical N-glycosylation site, we conducted LC-MS based peptide mapping analysis. When variable modification of G0F was set on each Asn in the sequence, two glycopeptides with Asn-X-Cys motif (XXXXXXXXXXN#K, named as GlycoPeP1 and XXXXXXXXXXXXXXXXN#SCXXXXXXXXXXXXXXSSGGGGSGG GGSGGGGSGGGGSXXXXXXXXXXX, named as GlycoPep2) were identified in BisP1 tryptic digests. Both glycopeptides were located in the VHH domain, consistent with the data from rCE-SDS . In contrast, the two glycopeptides were barely identified in BisP2 (data not show). MS and MS/MS spectra of GlycoPep2 eluted at 47.8 min are displayed in Figure 9. The oxonium ion-containing MS/MS spectra were considered as a vital evidence of glycosylation modification [26]. The presence of oxonium ions (m/z 126.055, m/z 138.055, m/z 144.065, m/z 186.076, m/z 204.086) with a series of b/y ions confirmed this noncanonical Asn-Ser-Cys sequence, instead of the Asn-X-Ser/Thr motif. For GlycoPep1, the tandem mass spectra are relatively poorer than GlyPep2, due to lower precursor ion intensity. Fragments peaks were assigned manually based on the MW of the glycans and peptide of GlycoPep1, as shown in Figure 10. Interestingly, the two glycopeptides behave very differently in tandem mass spectrometry. GlycoPep2 generates more b/y ions from peptide backbone fragmentation, while more ions representing the peptide backbone with glycan residue attached are observed in GlycoPep1. It is possible that the locations of N-glycosylation sites within the peptides and the length of the peptides have a combined impact on the fragmentation behaviour of glycopeptides. Nevertheless, two potential glycosites on the VHH domain were identified, both with a noncanonical Asn-X-Cys sequence.

To confirm the two atypical glycosites and how much each contributes to the impurity species, a full spectrum of analysis on proteins with one or two of the glycosites abolished were carried out. Firstly, we performed site-directed mutagenesis separately, and obtained glutamine substitution at the asparagines, namely N1Q (Asn-Lys-Cys site mutation) variant and N2Q (Asn-Ser-Cys site mutation) variant. Double-site random mutagenesis was then conducted, generating a mutant with aspartic acid substitution at Asn-Lys-Cys site and glycine substitution at Asn-Ser-Cys site, named as N1D/N2G. Mass spectrometry and rCE-SDS were first performed on the samples after initial protein. A capture to characterize the glycosylation profile. According to the intact mass deconvolution results (Figure 11), main peak showed a mass of 176681 Da for the wild type (WT), 176707 Da for two N1Q and N2Q mutants, and 176565 Da for N1D/N2G mutant, consistent with the theoretical MW of each construct. A mass shift of 1445 Da, matched with the G0F glycosylation modification, was detected in WT, N1Q and N2Q, but not in N1D/N2G. The extent of VHH glycosylation on each site can be reliably estimated by calculating the percentage of the shoulder peak in total HC (Figure 12). By abolishing both glycosites, N1D/N2G is the only construct with a single HC peak in the rCE-SDS analysis. N-glycosylation at Asn-Ser-Cys motif has a higher occupancy level (16.5 %) than that at Asn-Lys-Cys motif (5.5%), which is in accordance with the different precursor ion intensities of GlycoPep1 and GlycoPep2. Interestingly, the WT N-glycosylation occupancy is 19.1% in rCE-SDS analysis, which is almost equal to the sum of the N-glycosylation occupancy at each site.

Further comparison between WT and N1D/N2G were performed in nrCE-SDS analysis, Fd+VHH fragments MW determination and iCIEF analysis. As shown in Figure 13a, the nrCE-SDS main peak of WT was broader and asymmetrical, suggesting mixed populations with size differences. The main peak became sharp and symmetrical which indicated a more homogeneous species after the mutations abolishing the two atypical glycosylation sites. In the MW determination (Figure 13b), G0F modification was detected in Fd+VHH fragments of WT but not in N1D/N2G, confirming no glycans being added without the two Asn-X-Cys motifs. In iCIEF profiles (Figure 12c), N1D/N2G (8.2) has a lower pI than WT (8.5), probably due to the introduction of aspartic acid. The acidic variants decreased from 23.9% to 11.7% without the two glycosylation sites. Therefore, the N1D/N2G mutant has less charge heterogeneity. Our results confirmed that the N-glycosylation at two atypical Asn-X-Cys sites in the VHH domain of BisAb leads to increased size and charge heterogeneity.

Finally, the IEC purification was performed for N1D/N2G to answer the initial question why the BisAb has a two-peak elution profile. As expected, the double mutant elutes as a single peak in IEC purification (Figure 13d). In summary, our data demonstrated that the two-peak elution profile in the IEC purification of BisAb was caused by a rare N-glycosylation in the VHH domain. The extra neutral glycans, mainly G0F, may have changed the charge distribution of the BisAb, leading to a more acidic characteristic in the molecule, and the unique IEC profile.

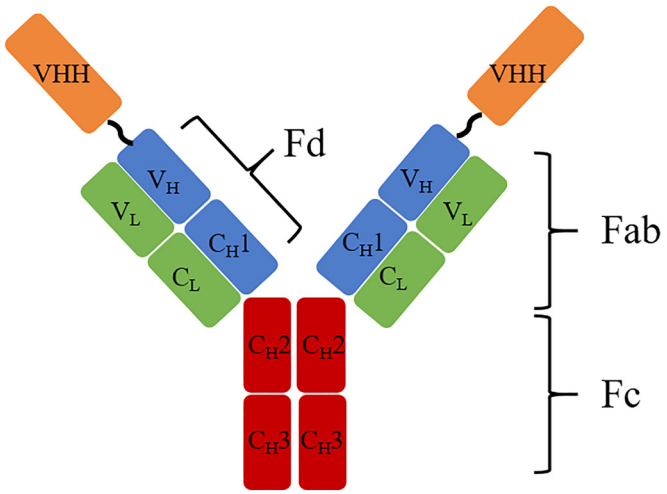

Figure 1: The domain structure of the BisAb. The domains of BisAb are labelled and coloured accordingly. The VHH domain (orange) is fused to the N terminus of the Fd region (blue) of a full-length IgG4 (light chain in green, Fc region in red) by a (G4S)4 linker (short black curve).

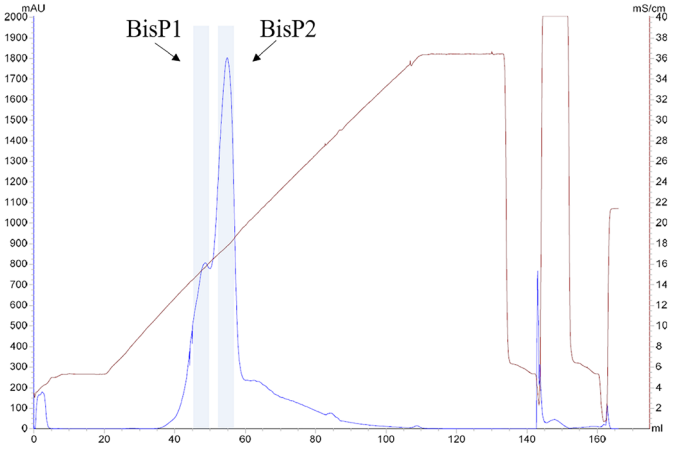

Figure 2: Double-peak elution behaviour of BisAb in IEC. Protein A column elution was loaded on an equilibrated column with Capto MMC Impres resin. A 0-400 mM NaCl linear gradient elution was performed at pH6.0. The brown line represents the conductivity. The blue line represents the A280 absorption. Blue shaded fractions were collected and pooled together as BisP1 and BisP2.

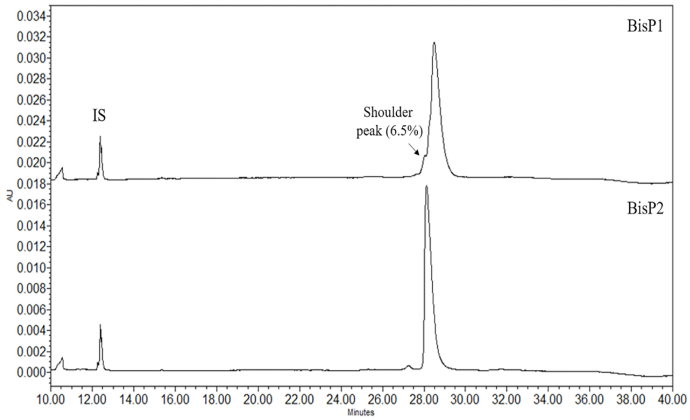

Figure 3: Non-reduced CE-SDS analysis for BisP1 and Bis P2. The main peak of BisP1 was asymmetrical with a migration time of 28.48 min, with a leading shoulder peak (6.5%). In comparison, BisP2 showed a typical nrCE-SDS profile with main peak elution time of 28.12 min. A 10-kDa protein was added in each sample as Internal Standard (IS) for migration time correction.

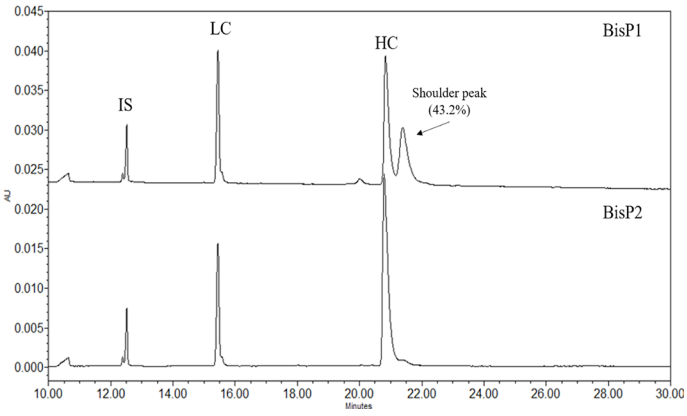

Figure 4: rCE-SDS analysis for BisP1 and Bis P2. The Light Chain (LC) for BisP1 and both LC and the Heavy Chain (HC) of BisP2 migrate at expected time (15.45 min for LC and 20.83 min for HC). Two peaks were detected for HC of BisP2, with the leading peak running at the expected time. The lagging peak, with migration time of 21.38 min, accounts for roughly 43.2% of total HC in BisP1. A 10-kDa Internal Standard (IS) was also included in the assay.

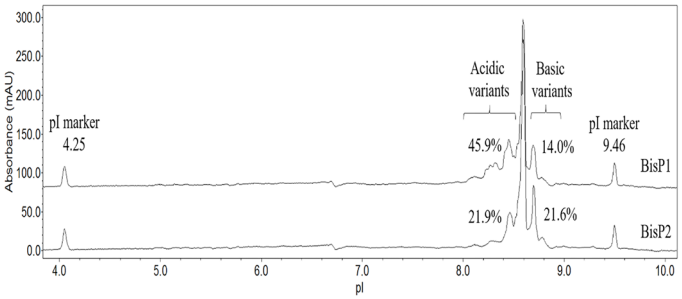

Figure 5: Imaged Capillary Isoelectric Focusing (iCIEF) profiles of BisP1, and BisP2.
Both BisP1 and BisP2 share the same isoelectric point (pI) of 8.5. The ratios of acidic and basic variants of each sample are labeled and BisP1, with enriched impurities, has more acidic and less basic populations than BisP2. Two markers (pI 4.25 and pI 9.46) were added in each sample for calibration of the isoelectric point.

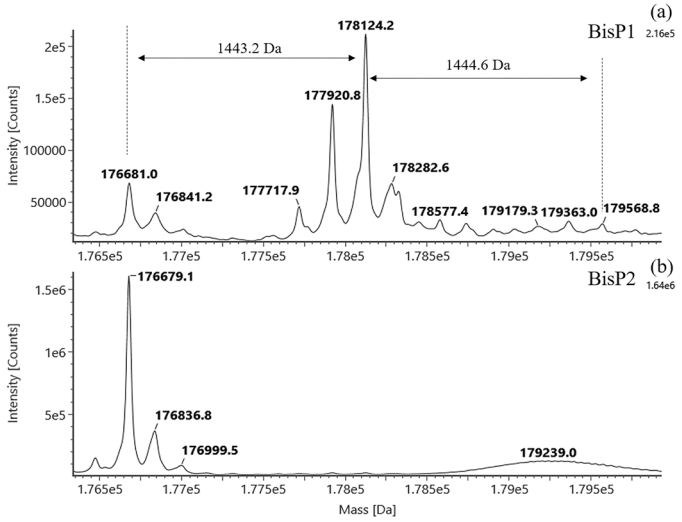

Figure 6: Deconvoluted mass spectra of BisP1 (a) and BisP2 (b). The observed MW for both BisP1 (176681 Da) and BisP2 (176679 Da) match the theoretical number (176682 Da). One and two of 1.44K Da-mass addition were observed in BisP1, which is consistent with the molecular weight of glycan G0F (1445.3 Da). Sequential mass difference of 203 Da from two neighbouring peaks was also detected in BisP1, indicating possible additions of acetyl hexosamine.

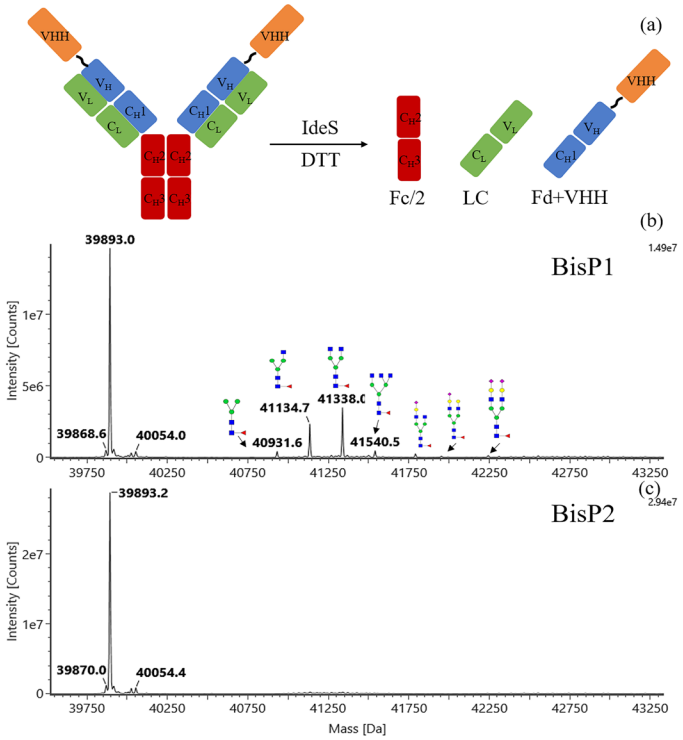

Figure 7: Deconvoluted mass spectra of BisP1 and BisP2 after DTT and IdeS treatment. (a) DTT and IdeS protease treatment generates Fc/2, LC and Fd+VHH fragments. The Molecular weight analysis was performed for all the fragments. Mass additions corresponding to potential glycosylation were found in the Fd+VHH of BisP1 (b), but not in BisP2 (c). Blue squares, green circles, yellow circles, magenta diamonds, and red triangles correspond to GlcNAc, Man, Gal, sialic acid, and Fuc, respectively.

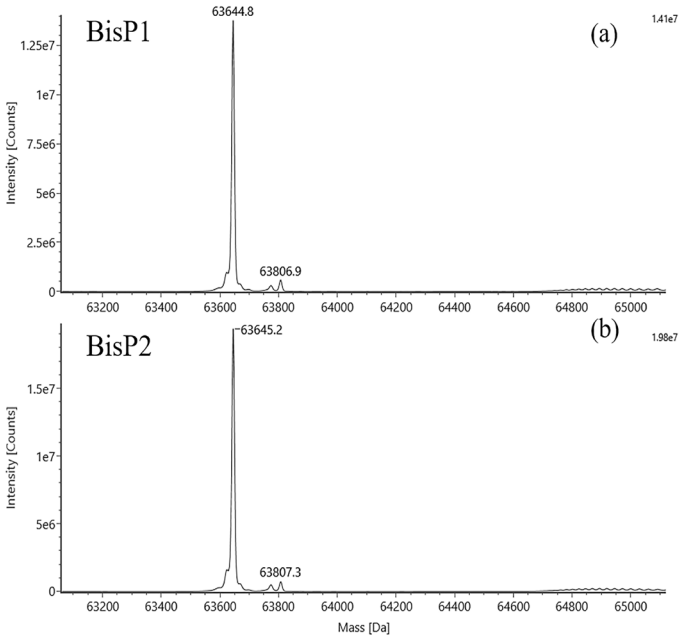

Figure 8: Deconvoluted mass spectra of deglycosylated heavy chain of BisP1 (a) and BisP2 (b). The observed HC MW for both BisP1 (63645 Da) and BisP2 (63645 Da) match the theoretical number (65044 Da). Neither 1445 Da mass addition nor +203 Da peak cluster can be detected.

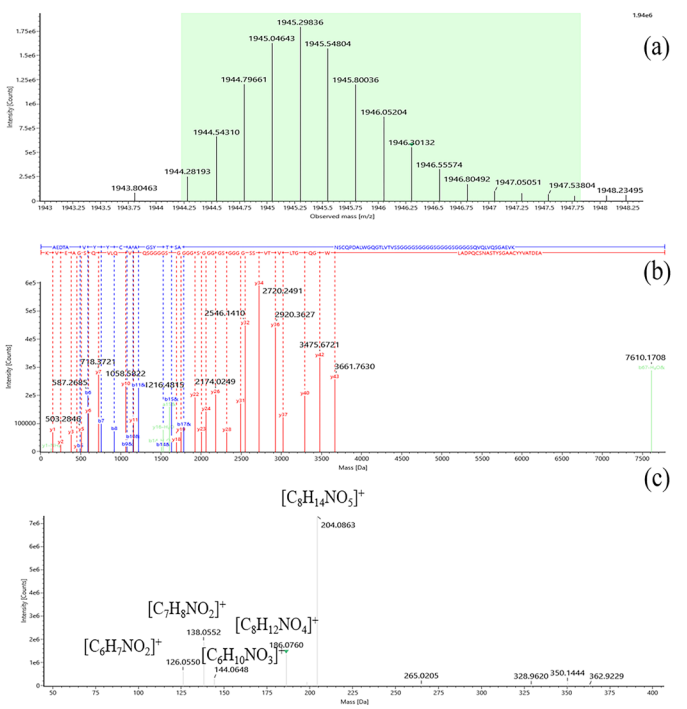

Figure 9: MS and MS/MS spectra of GlycoPep2. (a) Full scan mass spectrum of four-fold charged GlycoPep2. (b) Matching of the b and y ions from the amino acid sequence confirmed the identity of this peptide. (c) Matching of the oxonium ions confirmed the glycosylation modification.

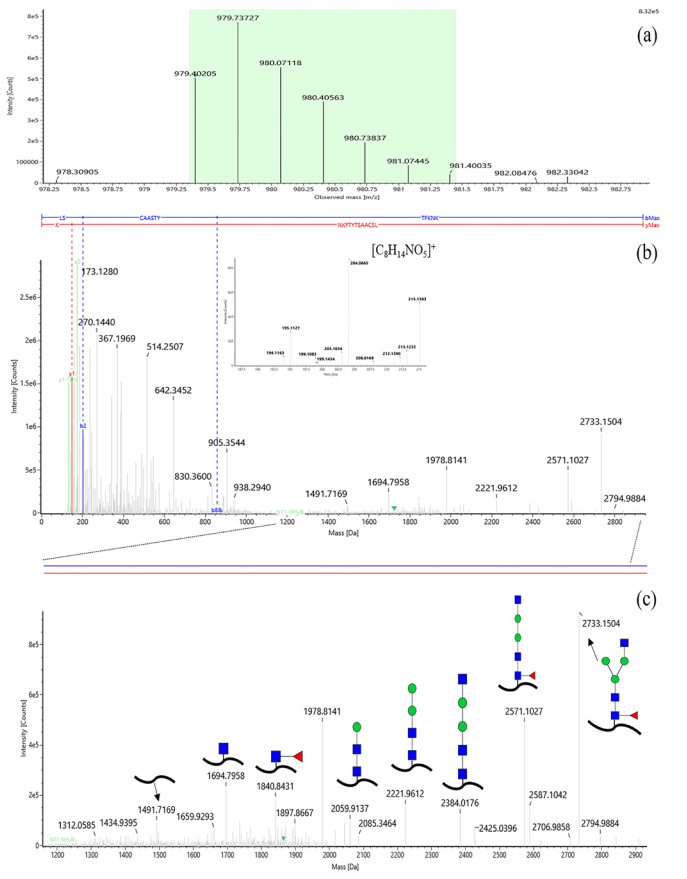

Figure 10: MS and MS/MS spectra of GlycoPep1. (a) Full scan mass spectrum of three-fold charged GlycoPep1. (b) Full scan tandem mass spectrum of GlycoPep1 and matched oxonium ion. (c) Matching of the fragments of GlycoPep1. Blue squares, green circles, red triangles correspond to GlcNAc, Man and Fuc, respectively.

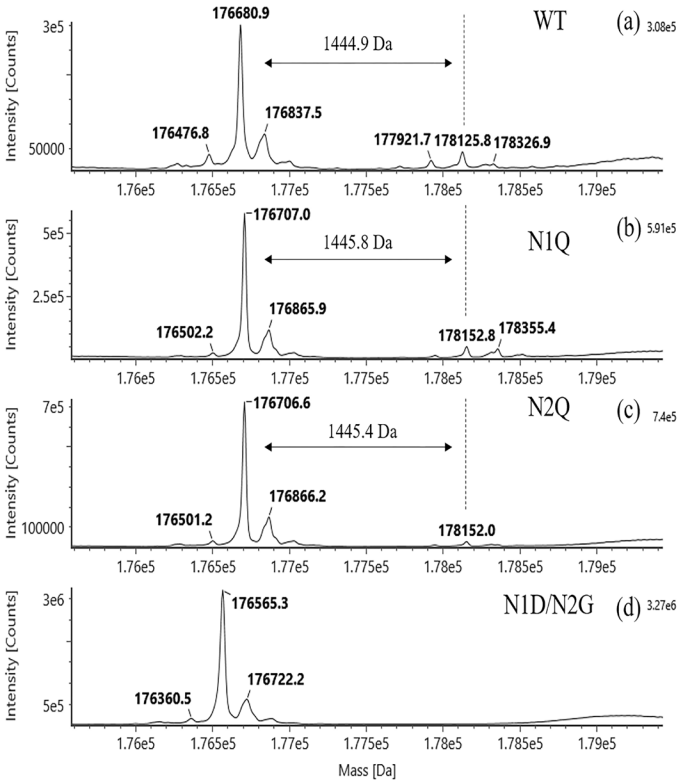

Figure 11: Molecular weight comparison between wild type and mutants. The comparison of deconvoluted mass spectra for protein A elution of Wild Type (WT) (a), N1Q (b), N2Q (c) and N1D/N2G (d). The observed MW of the main peak matches the theoretical MW well for all four constructs. 1.44 K Da-mass addition were observed in WT, N1Q and N2Q, but not in N1D/N2G.

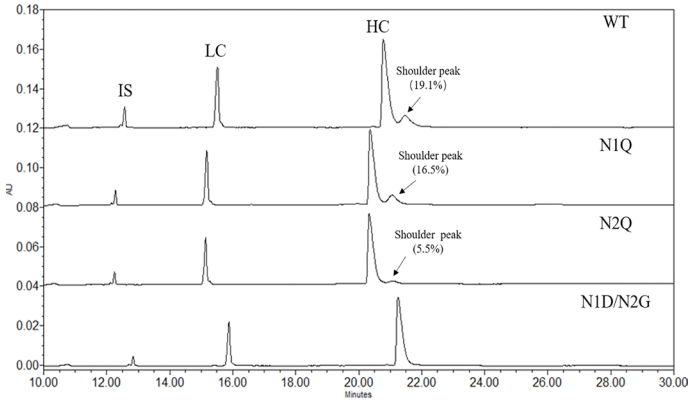

Figure 12: The rCE-SDS profiles for WT, N1Q, N2Q and N1D/N2G. Shoulder peaks on the right side of HC main peak with varying ratio were detected in WT (19.1%), N1Q (16.5%) and N2Q (5.5%). The double mutants N1D/N2G abolished the shoulder peak completely.

Discussion

Native mammalian proteins are often highly glycosylated, recombinant proteins for therapeutic purpose, such as Erythropoietin (rh-EPO), have a complicated glycan profile. In contrast, antibodies often have a relatively simple glycan profile. VHH is an 12-15 kDa antigen binding fragment, that is gaining popularity in therapeutic area. Small in size gives VHH unique advantages in designing multispecific antibodies and in therapeutics requiring better absorption or stability, such as inhaler or subcutaneous administration. However, the VHH glycosylation is much understudied. We first report and characterize the glycosylation profile of VHH domain in our protein of interest.

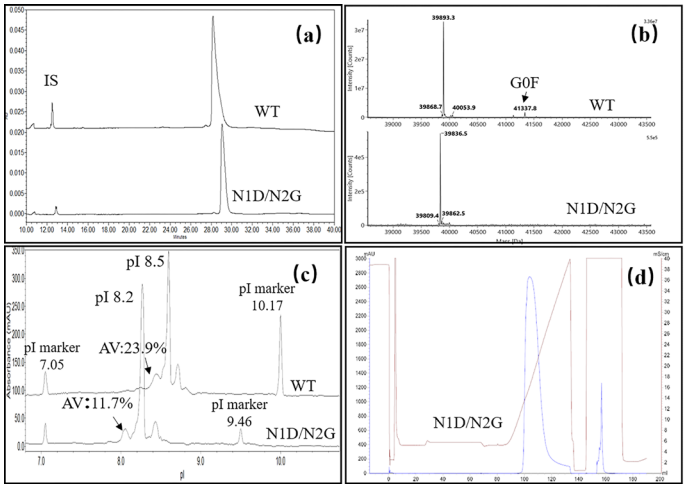

Figure 13: A comprehensive comparison between WT and the double mutant N1D/N2D. The nrCE-SDS (a), MW of Fd+VHH (b) and iCIEF (c) profiles were compared between the protein A elution of WT and of N1D/N2G. N1D/N2G mutant was further purified by IEC (d). (a) Broader and asymmetrical CE-SDS main peak was observed for WT, while N1D/N2G showed a single sharp peak. (b) G0F modification was detected in WT but not in N1D/N2G. (c) The pI changed from WT 8.5 to mutant 8.2 with N1D and N2G mutations. The Acidic Variants (AV) ratio reduced from 23.9% to 11.7% after the double mutation. (d) The mutant exhibited a single-peak elution behaviour as expected.

Glycosylation at the non-consensus Asn-X-Cys motif was reported as early as in 1990's in human plasma protein C [27]. Recombinant Human Epidermal Growth Factor Receptor produced by CHO cells was also found to be linked with oligosaccharide chain on Asn32-X33-Cys34 [28]. Hui Zhang's group conducted a comprehensive glycoproteomic analysis for CHO-K1 cell [29]. 1162 glycosites including 7 glycosites with Asn-X-Cys motif were identified. In our study, the discovery of rare glycosylation was originated from an abnormal IEC profile and the preliminary identification was conducted by CE-SDS analysis and MW analysis. Glycosites were localized by peptide mapping and further confirmed by mutagenesis.

Notably, at both Asn-X-Cys sites in the VHH domain of this BisAb, the glycan was predominately G0F. G0F, a complex bi-antennary-structured N-glycans with α1, 6-fucose linked to the chitobiose core (Figure 7b), is the most abundant type of glycan at the conserved Asn-297 site, located in the Fc domain of human IgGs. By contrast, N-glycosylation in the Fab domain of IgG molecules usually has higher level of sialyation and galactosylation [30,31]. The difference may come from factors involving glycosidase/glycosyltransferase activity, host cell line and manufacturing process.

According to the CE-SDS, the glycosylation occupancy in the VHH domain (16.5% for the Asn-Ser-Cys site and 5.5% for the Asn-Lys-Cys site) in our study was much lower than that in the Fc region (>95%) [32]. Although sulfhydryl group in cysteine can replace the usual hydroxyl group in serine or threonine as a hydrogen bond acceptor in the glycosylation reaction, the glycosylation efficiency was dramatically reduced [27]. The phenomenon was supported by Gerber et al using site specific mutation. Significant reduction in the in vitro affinity and decreased glycosylation rate were observed for acceptor peptides containing Asn-X-Cys sequon, comparing with those containing Asn-X-Ser/Asn-X-Thr motif [33]. This may partially explain the lower glycosylation occupancy for both sites identified, while local steric limitations may also weigh in.

Intriguingly, the additional neutral glycans in the VHH domain changed the charge behaviour of the protein: an increase in acidic peaks in iCIEF analysis. This explains why our BisAb elutes as two peaks during IEC purification, but the underlying mechanism remains unclear. Similar observations were reported by Valliere-Douglass [22,34]. Glycosylation is a complicated process that starts in the ER during protein translation and folding. It is possible that the addition of glycans changes the local conformation of the protein. Alternatively, the solvent accessibility of certain charged residues may be affected, resulting in a redistribution of surface charge. Another possibility lies in the steric effect of the extra glycans. Glycans, like G0F, is about 10 times bigger than a single amino acid. It could shield the charge of a spatially nearby amino acid. One of the glycosites in our study is Asn-Lys-Cys. The extra glycan on Asn is in proximity to the basic lysine. Shading the side chain of lysine could raise the acidity of the whole molecule. Nevertheless, it is an alarming effect that neutral glycan can change protein behaviour significantly, and requires more attention during process and method development.

Sequence optimization is important for the developability of a therapeutic candidate. The common modifications like oxidation, deamidation and glycosylation are usually highlighted and the hot spots containing these modifications are characterized carefully and avoided in variable domains when possible. Glycosylation in the variable domain dramatically increases the charge and size heterogeneity, due to the complex glycan composition and the varying glycosylation occupancy. The heterogeneity adds risks to variations among production batches and difficulties to the quality control, as well as characterization workload to any potential therapeutic candidate. Until now, most attentions are drawn to the consensus N-glycosylation sequence Asn-X-Ser/Asn-X-Thr. Our findings strongly suggest that other glycosylation sites like Asn-X-Cys should be carefully examined during early discovery phase.

Acknowledgements

The work is funded by Shanghai MabGen Biotech and Beijing Tuojie Biotech.

References

1. Sifniotis V, Cruz E, Eroglu B, Kayser V. Current advancements in addressing key challenges of therapeutic antibody design, manufacture, and formulation Antibodies. 2019; 8: 36.

2. Sharkey B, Pudi S, Moyer IW, Zhong L, Prinz B, et al. Purification of common light chain IgG-like bispecific antibodies using highly linear pH gradients MAbs. 2017; 9: 257-268.

3. Kimerer LK, Pabst TM, Hunter AK, Carta, G. Chromatographic behavior of bivalent bispecific antibodies on cation exchange columns. I. Experimental observations and phenomenological model. J Chromatogr A. 2019; 1601: 121-132.

4. Guo J, Zhang S, Carta, G. Unfolding and aggregation of a glycosylated monoclonal antibody on a cation exchange column. Part I. Chromatographic elution and batch adsorption behavior. J Chromatogr. A. 2014; 1356: 117-128.

5. Guo J, Carta G. Unfolding and aggregation of a glycosylated monoclonal antibody on a cation exchange column. Part II. Protein structure effects by hydrogen deuterium exchange mass spectrometry. J Chromatogr. A. 2014; 1356: 129-137.

6. Guo J, Creasy AD, Barker G, Carta G. Surface induced three-peak elution behavior of a monoclonal antibody during cation exchange chromatography. J Chromatogr. A. 2016; 1474: 85-94.

7. Wu H, Truncali K, Ritchie J, Kroe-Barrett R, Singh S, et al. Roberts CJ. Weak protein interactions and pH- and temperature-dependent aggregation of human Fc1. MAbs. 2015; 7: 1072-1083.

8. De Leoz MLA, Duewer DL, Fung A, Liu L, Yau HK, et al. NIST Interlaboratory Study on Glycosylation Analysis of Monoclonal Antibodies: Comparison of Results from Diverse Analytical Methods. Mol. Cell. Proteomics. 2020; 19: 11-30.

9. Cymer F, Beck H, Rohde A, Reusch D. Therapeutic monoclonal antibody N-glycosylation-Structure, function and therapeutic potential. Biologicals. 2018; 52: 1-11.

10. Ehret J, Zimmermann M, Eichhorn T, Zimmer A. Impact of cell culture media additives on IgG glycosylation produced in Chinese hamster ovary cells. Biotechnol Bioeng. 2019; 116: 816-840.

11. Goh JB, Ng SK. Impact of host cell line choice on glycan profile. Critical Reviews in Biotechnology. 2018; 38: 851-867.

12. Zhang P, Woen S, Wang T, Liau B, Zhao S, et al. Challenges of glycosylation analysis and control: an integrated approach to producing optimal and consistent therapeutic drugs. Drug Discovery Today. 2016; 21: 740-765.

13. Zhu L, Guo Q, Guo H, Liu T, Zheng Y, et al. Versatile characterization of glycosylation modification in CTLA4-Ig fusion proteins by liquid chromatography-mass spectrometry. MAbs. 2014; 6: 1474-1485.

14. Houel S, Hilliard M, Yu YQ, McLoughlin N, Martin SM, et al. N- and O-glycosylation analysis of etanercept using liquid chromatography and quadrupole time-of-flight mass spectrometry equipped with electron-transfer dissociation functionality Anal Chem. 2013; 86: 576-584.

15. Cho IH, Lee N, Song D, Jung SY, Bou-Assaf G, et al. Evaluation of the structural, physicochemical, and biological characteristics of SB4, a biosimilar of etanercept. MAbs.2016; 8: 1136-1155.

16. Ohyama Y, Nakajima K, Renfrow MB, Novak J, Takahashi K. Mass spectrometry for the identification and analysis of highly complex glycosylation of therapeutic or pathogenic proteins. Expert Rev. Proteomics. 2020; 17: 275-296.

17. Zielinska DF, Gnad F, Wisniewski JR, Mann M. Precision mapping of an in vivo N-glycoproteome reveals rigid topological and sequence constraints. Cell. 2010; 141: 897-907.

18. Sun SS, Zhang H. Identification and validation of atypical N-glycosylation sites. Anal Chem. 2015; 87: 11948-11951.

19. Sun SS, Hu Y, Jia L, Eshghi ST, Liu Y, et al. Site-specific profiling of serum glycoproteins using N-linked glycan and glycosite analysis revealing atypical N-glycosylation sites on Albumin and alpha-1B-Glycoprotein. Anal Chem. 2018; 90: 6292-6299.

20. Yasuda D, Imura Y, Ishii S, Shimizu T, Nakamura M. The atypical N-glycosylation motif, Asn-Cys-Cys, in human GPR109A is required for normal cell surface expression and intracellular signaling. FASEB J. 2015; 29: 2412-2422.

21. Lowenthal MS, Davis KS, Formolo T, Kilpatrick LE, Phinney KW. Identification of novel N-glycosylation sites at non-canonical protein consensus motifs. J Proteome Res. 2016; 15: 2087-2101.

22. Valliere-Douglass JF, Kodama P, Mujacic M, Brady LJ, Wang W, et al. Asparagine-linked oligosaccharides present on a non-consensus amino acid sequence in the CH1 domain of human antibodies. J Biol Chem. 2009; 284: 32493-32506.

23. Valliere-Douglass JF, Eakin CM, Wallace A, Ketchem RR, Wang W, et al. Glutamine-linked and non-consensus asparagine-linked oligosaccharides present in human recombinant antibodies de-

fine novel protein glycosylation motifs J Biol Chem. 2010; 285: 16012-16022.

24. Spahr CS, Daris ME, Graham KC, Soriano BD, Stevens JL, et al. Discovery, characterization, and remediation of a C-terminal Fc-extension in proteins expressed in CHO cells. MAbs. 2018; 10: 1291-1300.

25. Kotia RB, Raghani AR. Analysis of monoclonal antibody product heterogeneity resulting from alternate cleavage sites of signal peptide. Anal Biochem. 2010; 399: 190-195.

26. Madsen JA, Farutin V, Lin YY, Smith S, Capila, I. Data-independent oxonium ion profiling of multi-glycosylated biotherapeutics. MAbs.2018; 10: 968-978.

27. Miletich JP, Broze Jr GJ. Beta protein C is not glycosylated at asparagine 329. The rate of translation may influence the frequency of usage at asparagine-X-cysteine sites. J Biol Chem. 1990; 265: 11397-11404.

28. Sato C, Kim JH, Abe Y, Saito K, Yokoyama S, et al. Characterization of the N-oligosaccharides attached to the atypical Asn-X-Cys sequence of recombinant human epidermal growth factor receptor. J Biochem. 2000; 127: 65-72.

29. Yang G, Hu Y, Sun S, Ouyang C, Yang W, et al. Comprehensive Glycoproteomic analysis of chinese hamster ovary cells Anal Chem. 2018; 90: 14294-14302.

30. Bondt A, Rombouts Y, Selman MHJ, Hensbergen PJ, Reiding KR, et al. Immunoglobulin G (IgG) Fab glycosylation analysis using a new mass spectrometric high-throughput profiling method reveals pregnancy-associated changes. Mol. Cell. Proteomics.2014; 13: 3029-3039.

31. van de Bovenkamp FS, Hafkenscheid L, Rispens T, Rombouts Y. The Emerging importance of IgG Fab glycosylation in immunity. J Immunol. 2016; 196: 1435-1441.

32. Cao C, Yu L, Zhang X, Dong X, Yuan J, et al. Calibration for quantitative Fc-glycosylation analysis of therapeutic IgG1-type monoclonal antibodies by using glycopeptide standards. Anal Chim. Acta. 2021; 1154: 338306.

33. Gerber S, Lizak C, Michaud G, Bucher M, Darbre T, et al. Mechanism of Bacterial Oligosaccharyltransferase: In vitro quantification of sequon binding and catalysis. J Biol Chem. 2013; 288: 8849-8861.

34. Valliere-Douglass JF, Brady LJ, Farnsworth C, Pace, D, Balland, A, et al. O-Fucosylation of an antibody light chain: Characterization of a modification occurring on an IgG1 molecule. Glycobiology. 2009; 19: 144-152.