The Big Protein Project: Development and Assessment of Modern Protein Glycosylation Characterization Techniques

Michael Karfunkle[1*], Joshua Shipman[1], Kang Chen[2], Hongbin Zhu[1], You Zhuo[3]; Milani Patabandige[3], Di Wu[2], Mercy Oyugi[4], Sarah Rogstad[5*]

1. CDER/OPQ/OTR-Saint Louis, MO; 2. CDER/OPQ/OTR – White Oak; 3; CDER/OPQ/OBP; 4. CDRH/OSEL; 5. CDER/OPQ/IO; *Primary Contact

Abstract

Analytical characterization of the structure and chemistry of monoclonal antibodies (mAb) is critical for quality assessment of these products. In particular, glycosylation – the enzymatic attachment of oligosaccharides to the protein backbone – is a critical quality attribute (CQA) that must be well characterized and controlled before regulatory approval because different glycans can impact the safety and efficacy of mAbs and other therapeutic proteins. Glycans can differ in monosaccharide composition, branching, and connectivity, which makes analysis challenging. The Big Protein Project (BPP) was initiated with the aim of developing a panel of parallel analytical procedures to thoroughly characterize therapeutic protein drug products.

Analytical characterization was performed on multiple lots of rituximab from two sources, one approved (Firm A) and one not approved (Firm B) for the US marketplace by the FDA. Multiple orthogonal analytical techniques were performed including analysis of released glycans by hydrophilic interaction chromatography (HILIC) with fluorescent detection (FLD), liquid chromatography (LC)-Mass Spectrometry (MS)-based multi-attribute method (MAM), intact-mass LC-MS, and nuclear magnetic resonance (NMR) spectroscopy. Individual glycoforms detected by MS and HILIC-FLD were compared directly while NMR quantified *classes* of glycans (afucosylation, galactosylation and high-mannose). While each method possessed advantages and disadvantages, there was concurrence between methods in the observed lot-to-lot and manufacturer-based differences.

Introduction

N-glycans – oligosaccharides connected to asparagine (N) residues – can differ in saccharide composition (Table 1, Figure 1), branching, and connectivity which makes analysis challenging. The conventional workflow for N-glycan analysis is known as released glycan analysis. In this process the glycans are enzymatically released from the protein and the reducing end aldehyde is then labeled with a fluorophore. Glycans can then be separated and analyzed by HILIC-FLD.

Glycans can be analyzed via high resolution mass spectrometry (HRMS) using a variety of workflows at the intact, peptide, or glycan level. Modern mass spectrometers offering resolutions greater than 100,000 in combination with deconvolution software can be used to analyze fully intact or reduced proteins. Proteolytic digestion provides site-specific information by analyzing mAbs at the peptide level which also allows for the characterization of other CQAs such as oxidation and deamidation. Glycopeptide analysis is one aspect of MAM that had been increasingly implemented in industry Quality by Design and quality control roles. In addition, a middle-down NMR method was used to profile glycan distributions. These types of analyses offer minimal sample preparation compared to the other approaches discussed.

D-Galactose	G	○
N-acetyl-glucosamine	GN	□
D-Mannose	M	○
L-Fucose	F	△
N-acetyl-neuraminic acid	S	◇

Table 1. Symbols of monosaccharides present in rituximab *N*-glycans.

Materials and Methods

The BPP encompasses multiple methods including the conventional HILIC-FLD glycan analysis, 1- and 2- dimensional NMR, and mass spectrometry methods: MAM developed in-house, reduced-mass LC-MS, released glycan analysis by QTOF LC-MS and intact mass analysis. Due to space limitations, this poster does not include methods instructions in detail, but the primary contacts can provide them.

Multi-Attribute-Method (MAM)

Samples were prepared using a tryptic digest followed by LC-MS/MS on a Thermo Q-Exactive hybrid quadrupole-orbitrap MS with heated electrospray ionization source, 5 µg per replicate, 95 minute runtime.

Intact Mass LC-MS

Samples were buffer exchanged into 95:5 water:acetonitrile (ACN) with 0.1% formic acid using three wash/centrifugation cycles in 10 kDa molecular weight cutoff filters, 20 µg sample per replicate. Samples were analyzed using a 20 minute gradient on a Thermo Orbitrap Fusion operating at a resolution of 175,000.

Reduced Intact (RI) Mass LC-MS

Identical sample preparation to Intact Mass with the addition of tris(2-carboxyethyl)phosphine (TCEP) to the buffer exchange solution, 20 µg sample per replicate. Samples were analyzed using a 20 minute gradient on a Thermo Orbitrap Fusion operating at a resolution of 175,000.

Released Glycans by HILIC-FLD

Sample preparation included formic acid digestion followed by centrifuge filtration, deglycosylation with PNGase F, labeling with 2-AB, washing to remove excess labeling reagent, then analysis by HILIC-FLD. 200 µg per replicate, 47.5 minute runtime.

Nuclear Magnetic Resonance

Includes ^{1}H 1D NMR with solvent suppression and ^{1}H-^{13}C Heteronuclear Single-Quantum Coherence (HSQC), performed on both a 600 MHz magnet with nitrogen-based cryoprobe and on a 850 MHz magnet with a helium-based cryoprobe. 5000 µg per replicate, 10.5 hour runtime.

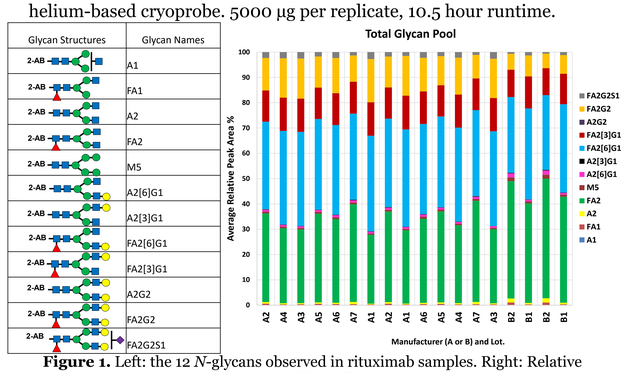

Figure 1. Left: the 12 *N*-glycans observed in rituximab samples. Right: Relative distributions of these glycans observed by HILIC-FLD. Each lot was analyzed in duplicate. The 4 most abundant *N*-glycans are classified as the major glycans: FA2G2, FA2[3]G1, FA2[6]G1, and FA2. The other 8 are classified as minor glycans.

Results and Discussion

HILIC-FLD analysis of released glycans is the conventional method for characterizing *N*-glycans in protein drugs and was used as the benchmark for the BPP. The glycan profiles of 9 lots of rituximab (7 Firm A and 2 Firm B) were determined by each method (Figure 2). The observable glycan CQAs differed across the various methods as presented in Figure 3.

Results and Discussion

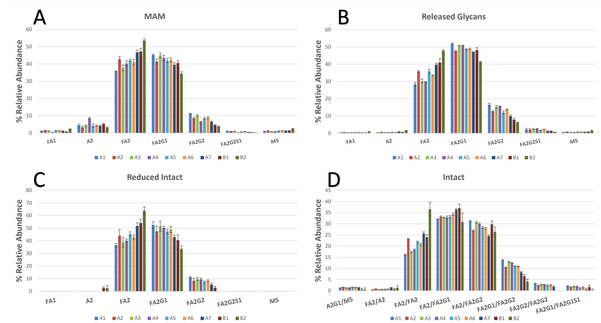

Figure 2. Comparison of the % relative abundance (%RA) of seven *N*-glycans in rituximab as observed by MAM (A), HILIC-FLD (B), Reduced intact MS (C), and Intact Mass MS (D), separated by lot. Error bars represent standard deviation.

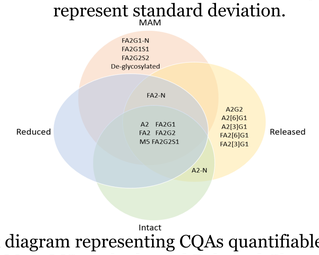

Figure 3. Venn diagram representing CQAs quantifiable by MAM, Intact Mass, Reduced-Mass MS methods, and Released Glycan by HILIC-FLD.

NMR: NMR samples can provide quantitative data on CQAs including total galactosylation, total fucosylation, and residual solvents, and qualitative information including identity by the anomeric "fingerprint" (Figure 4, Table 2), the presence of minor glycans, and excipients. However, due to the nature of NMR detection, the granularity of information it can provide is limited compared to MS, and requires orders of magnitude more sample.

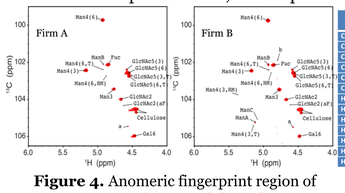

Figure 4. Anomeric fingerprint region of HSQC spectra of Firm A (left) and Firm B (right).

	GN1	GN2	Man3	Fuc3'	Man4	Man4'	GN5/5'a	GN5/5'b	Gal6/6'
C1	81.16	104.0	103.5	102.1	102.4	99.7	102.5	102.7	106.0
C2	57.59	57.66	72.98	71.04	79.36	79.49	57.61	58.09	73.75
C3	75.38	74.67	83.56	72.26	72.33	72.39	74.79	76.17	75.42
C4	81.28	82.59	68.29	74.98	70.16	70.17	81.73	72.75	71.42
C5	77.82	77.16	77.16	69.68	76.50	75.72	77.52	78.72	78.22
C6	69.49	62.75	68.44	-	64.52	64.32	62.77	-	63.42
H1	5.08	4.66	4.76	4.83	5.16	4.92	4.55	4.54	4.47
H2	3.80	3.76	4.28	3.81	4.21	4.10	3.74	3.74	3.59
H3	3.76	3.71	3.79	3.92	3.91	3.86	3.74	3.59	3.68
H4	3.84	3.73	3.86	3.79	3.52	3.52	3.72	3.47	3.94
H5	3.60	3.60	3.65	4.16	3.79	3.61	3.59	3.46	3.73
H6a	3.91	3.85	3.98	-	3.96	3.91	3.98	-	3.93
H6b	-	-	3.79	-	3.63	3.66	-	-	3.77

Table 2. ^{1}H-^{13}C HSQC chemical shift assignments for major glycans of rituximab. C and H numbers refer to the C or H nucleus position on the monosaccharide.

Comparability and Orthogonality of Methods

The accuracy and precision of these methods can be assessed by comparison of the results themselves, using the released glycans by HILIC-FLD as a benchmark (Figures 5 and 7), and by comparison of the percent coefficients of variation (%CV, Figure 6). There is a general trend that the lower the abundance of the glycan, the lower the accuracy and precision, which is to be expected. Of the three MS methods (MAM, intact mass, and reduced intact mass), intact mass demonstrated highest accuracy at low abundance but poorer precision than MAM or the non-MS benchmark HILIC-FLD.

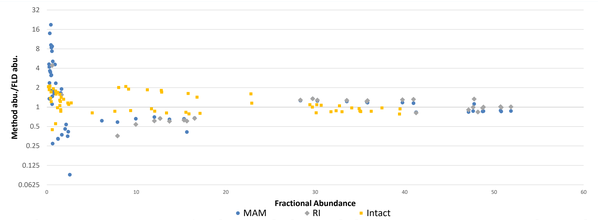

Figure 5. Accuracy: Ratio of method-observed abundance to HILIC-FLD-observed abundance vs. fractional abundance of glycan CQAs across LC-MS methods: MAM (blue), reduced-mass (grey), and intact-mass (yellow).

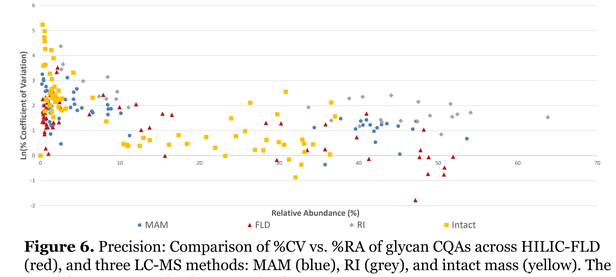

Figure 6. Precision: Comparison of %CV vs. %RA of glycan CQAs across HILIC-FLD (red), and three LC-MS methods: MAM (blue), RI (grey), and intact mass (yellow). The Ln of the %CV is used to linearize the data.

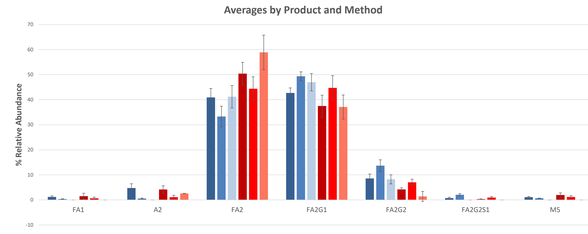

Figure 7. Comparison of results (%RA) separated by method and product. Error bars represent standard deviation. Note that the %RA is normalized to a total of 100% and the observable CQAs differ across methods (See Figure 3).

Conclusion

- Results demonstrate that agreement can be found between conventional and state-of-the-art analytical methods and that these high-resolution methods can provide increased confidence in comparative analytical studies, providing flexibility in method selection.
- Significant differences in major and minor *N*-glycan profiles were observed both between lots from the same manufacturer and between manufacturers. Tracking lot-to-lot variation is a critical part of product surveillance, and this data can be used to define acceptable bounds for these CQAs.
- Each method was validated according to ICH guidelines and their performance compared to HILIC-FLD demonstrates their relative advantages in terms of sensitivity, selectivity, time, and material costs and their mutual orthogonality.
- Future work includes assessing and developing similarly modern, high-resolution analytical methods for other protein drug CQAs such as higher order structure, aggregates, and impurities.