Facilitating released N-glycan identification using LC-MS and an extensive glycan library

Featuring SCIEX OS software

Zhiliang Xiao,[1] Ji Luo,[1] Lihai Guo,[1] Zhengwei Chen[2]
[1]SCIEX, China; [2]SCIEX, USA

In this technical note, an N-glycan library is introduced as a useful tool for fast identification of the released N-glycans from protein therapeutics. This library is fully integrated into SCIEX OS software for routine automated analysis. The three most commonly used labeling reagents—2-aminobenzamide (2-AB), procainamide (ProA) and RapiFluor (RF)—are incorporated to meet the diverse needs of biopharma-focused LC-MS analysis.

Biotherapeutics have been the fastest-growing modality in the therapeutic segment over the past decade. They have played an essential role in treating various diseases such as tumors, metabolic disorders and autoimmune diseases. Glycosylation is a common post-translational modification (PTM) in biotherapeutics: more than 70% of protein therapeutics are glycosylated proteins. Glycosylation has been shown to have a profound impact on the safety and efficacy of protein therapeutics and thus is usually considered to be a critical quality attribute (CQA). The high heterogeneity of N-glycosylation in protein therapeutics expressed in Chinese hamster ovary (CHO) cell lines brings significant challenges for the released glycan analysis. However, there are inherent rules for the N-glycosylation biosynthesis pathway in CHO cell lines. That is, every N-glycan starts with the formation of 2 GlcNAc and 3 Man at the core and is later modified by adding other sugar units to generate high-mannose, complex and hybrid N-glycans. The recognition of these conserved biosynthesis pathways could be used to predict and construct an N-glycan library.

An N-glycan library of 236 N-glycans was created based on the N-glycan biosynthesis pathway in CHO cell lines. Released glycan analysis of rituximab, a chimeric monoclonal antibody (mAb), is demonstrated using the TripleTOF 5600+ system from SCIEX. A single software platform (SCIEX OS software) was used, which offers a robust data processing solution for released glycan analysis, covering extracted ion chromatogram (XIC) generation, glycan identification, data filtering and report generation. This library can be applied to any SCIEX accurate mass system due to their similarity in fragmentation. The workflow is also applicable to other types of protein therapeutics.

Key features of the N-glycan library in SCIEX OS software

- A comprehensive picture of the N-glycans expressed in CHO cell lines is covered through an N-glycan library of 236 N-glycan structures corresponding to a total of 177 compositions.
- Fast processing and comprehensive identification of N-glycosylations of protein therapeutics can be achieved easily by applying the library in SCIEX OS software.
- Flexibility of the sample preparation is encompassed by incorporating all three of the most commonly used N-glycan labeling reagents (2-AB, ProA and RF) into the released N-glycan library presented.
- Versatility of the data acquisition platform is enabled by the compatibility of SCIEX OS software with processing data from all SCIEX accurate mass systems.

Methods

Sample preparation: The sample preparation was conducted according to the RF kit protocol. Briefly, 15 µg of the rituximab mAb were denatured at 95°C after being dissolved in RapiGest SF buffer. N-glycans were released by PNGase F and then labeled by RF The labeled N-glycans were purified using HILIC-SPE before being injected for LC-MS analysis.

LC-MS conditions: The processed samples were subject to LC-MS/MS analysis using the ExionLC AD system coupled to a TripleTOF 5600+ system in positive ionization mode. LC separation was achieved on a C18 column (ACQUITY UPLC Glycan BEH Amide, 1.7 µm, 2.1 x 150 mm) with 50 mM ammonium formate at pH 4 as mobile phase A and acetonitrile as mobile phase B at a flow rate of 0.2–0.4 mL/min. The column was kept at 60°C, and 10 µL of sample were injected. The experimental details are summarized in Table 1 and Table 2.

Table 1. MS parameters.

Parameter	MS	MS/MS
Scan mode	*TOF-MS*	*IDA dependent*
Gas 1	*50 psi*	
Gas 2	*50 psi*	
Curtain gas	*25 psi*	
Source temperature	*450°C*	
Ion spray voltage	*500 V*	
Declustering potential	*20 V*	
Collision energy	*10 V*	*20 V*
CAD gas	*7*	
Maximum candidate ions	*5*	
Intensity threshold	*300 cps*	
Charge states	*2 to 5*	
Exclusion time	*1 s after 2 occurrences*	
Start mass	*200 m/z*	*100 m/z*
Stop mass	*2,000 m/z*	*3,000 m/z*
Accumulation time	*0.25 s*	*0.1 s*
Time bins to sum	*4*	*4*

Table 2. Gradient used for LC separation.

Time [min]	Mobile phase A [%]	Mobile phase B [%]
0	*25*	*75*
35	*40*	*60*
36.5	*100*	*0*
39.5	*100*	*0*
43.1	*25*	*75*
47.6	*25*	*75*
55	*25*	*75*

Data processing: The constructed N-glycan library was implemented into the SCIEX OS software (Figure 2). The embedded “Quantitation and targeted identification” workflow was used to allow relative quantification and targeted identification of various N-glycan species. In the workflow, the accurate mass and chemical formula of each N-glycan from the library at different charge states (+2, +3) were entered in the “Components” section in the software. Generic integration parameters were set for peak integration, which can be adjusted in the results file if needed. The processing method was then used to reduce the raw data files generated from the released N-glycan samples.

Results and discussion

Although the structures of N-glycan are diverse with different compositions and branches, the biosynthesis pathway follows a specific conservative pattern. The common core Man3GlcNAc2 is formed first, and complex N-glycans are then generated by the addition of further sugar units to this core structure. High-mannose N-glycans are created by the addition of up to six mannoses. In complex N-glycans, the “antennae” initiated by GlcNAc extend the core and can be further extended by galactose, sialic acids, etc. In hybrid N-glycans, mannoses extend one arm of the core, and one or two GlcNAcs extend the other arm.

spectra (Figure 2). Flexible filtering options for the results table—based on user-defined criteria such as mass accuracy, isotopic distribution, retention time and integration area—greatly facilitate the confirmation process within SCIEX OS software (area A in Figure 2). During review of the XICs panel, integration parameters can be adjusted to ensure reliable integration accuracy and accurate relative quantification of each N-glycan species for the chosen LC settings. For each N-glycan identification, the MS/MS spectra information was displayed alongside the MS spectra and retention times, which allows results table as well as the XICs, MS spectra and MS/MS further structural elucidation by looking into signature ions for isomers (areas B, C and D in Figure 2).

Oxford Notation	Formula	Mono Mass (Da, free end)	CFG Structure	2-AB		ProA		RF	
M3	C34H56O25N2	910.3278		C41H64O26N4	1028.3809	C47H77O26N5	1127.4857	C51H79O27N7	1221.5024
M4	C40H66O30N2	1072.3086		C47H74O31N4	1190.3617	C53H87O31N5	1289.4665	C57H89O32N7	1383.4832
M5	C46H76O35N2	1234.4334		C53H84O36N4	1352.4865	C59H97O36N5	1451.5913	C63H99O37N7	1545.6080
M6	C52H86O40N2	1396.4863		C59H94O41N4	1514.5394	C65H107O41N5	1613.6442	C69H109O42N7	1707.6609
M7	C58H96O45N2	1558.5391		C65H104O41N4	1676.5922	C71H117O46N5	1775.6970	C75H119O47N7	1869.7137
M8	C64H106O50N2	1720.5919		C71H114O51N4	1838.6450	C77H127O51N5	1937.7498	C81H129O52N7	2031.7665
M9	C70H116O55N2	1882.6447		C77H124O56N4	2000.6978	C83H137O56N5	2099.8026	C87H139O57N7	2193.8193
A1	C42H69O30N3	1113.4072		C49H77O31N5	1231.4603	C55H90O31N6	1330.5651	C59H92O32N8	1424.5818
A1F	C48H79O34N3	1259.4651		C55H87O35N5	1377.5182	C61H100O35N6	1476.6230	C65H102O36N8	1570.6397
A1G1	C48H79O35N3	1275.4600		C55H87O36N5	1393.5131	C61H100O36N6	1492.6179	C65H102O37N8	1586.6346

Figure 1. Overview of a part of the N-glycan library. The N-glycan library consists of 236 entries. Oxford notation was used to name the N-glycans, while CFG notation was used for the structure representation. Chemical formulas and monoisotopic masses from both native N-glycans and labeled N-glycans (2-AB, ProA and RF) were included in the library.

Based on the structural nature of N-glycans, described previously, the N-glycan library was constructed with a total of 236 N-glycan structures. To be consistent with the N-glycan naming system used in the glycobiology community, the Oxford naming nomenclature was adopted for the name annotation, while the Consortium for Functional Glycomics (CFG) nomenclature was used for the structural annotation. To provide a comprehensive N-glycan library that accounts for different labeling reagents, the three most commonly used N-glycan labeling reagents (2-AB, ProA and RF) were incorporated into the library with their chemical formula and accurate mass information (Figure 1).

The constructed library was incorporated into the processing method in SCIEX OS software, which was used for processing raw data in batches and to generate results files. Subsequently, a user can validate the N-glycan identifications based on the

By using this approach for the released N-glycan analysis of rituximab, a total of 16 N-glycans were confidently identified at various levels, as shown in Figure 3. The report generated in SCIEX OS software provided information such as glycan composition, accurate mass, integration area and chemical formula (Figure 4). This report can also be customized for different characterization requirements.

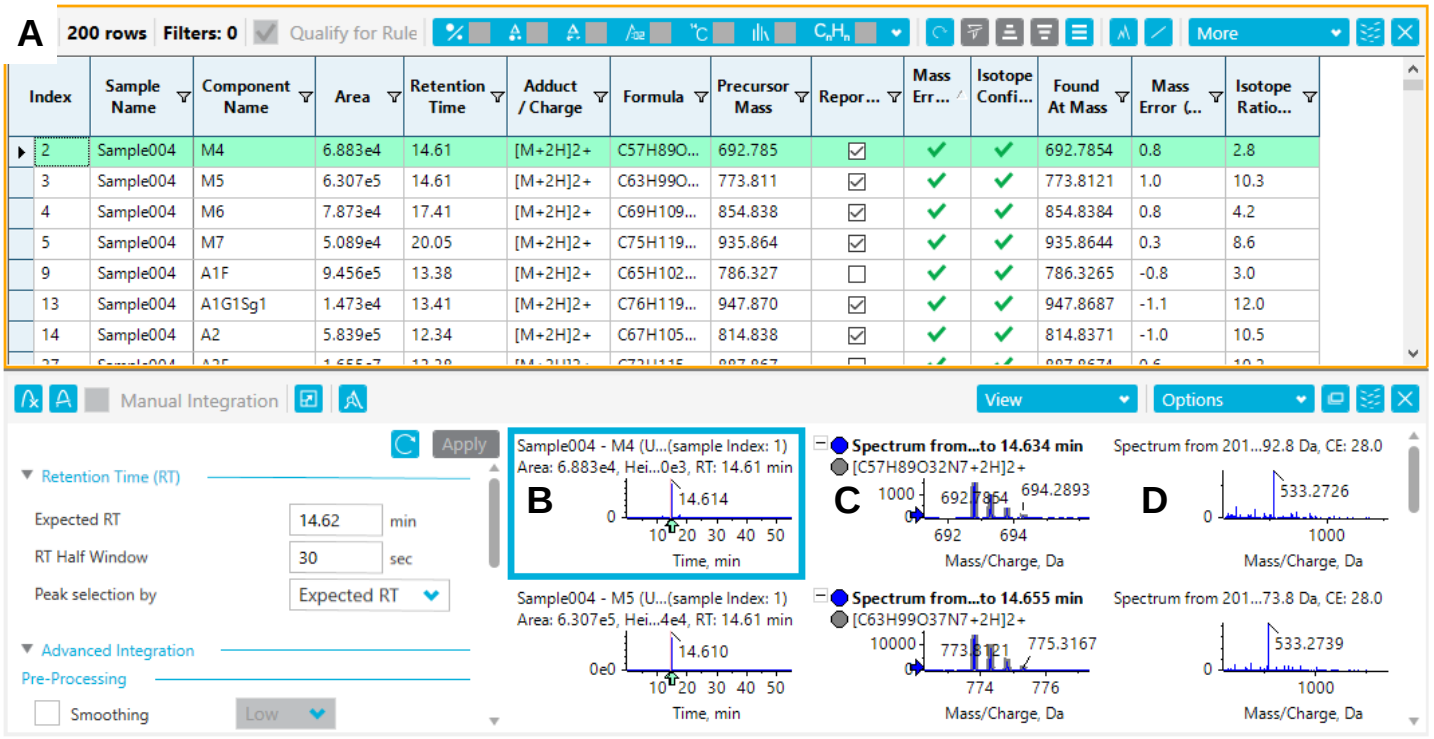

Figure 2. Data processing in SCIEX OS software. Based on the accurate mass match, the N-glycan composition was assigned to each peak (A). The assigned N-glycans were further validated by checking the XIC integration area and retention time (B) as well as the isotopic distribution of the precursor (C) and the derived MS/MS spectrum (D).

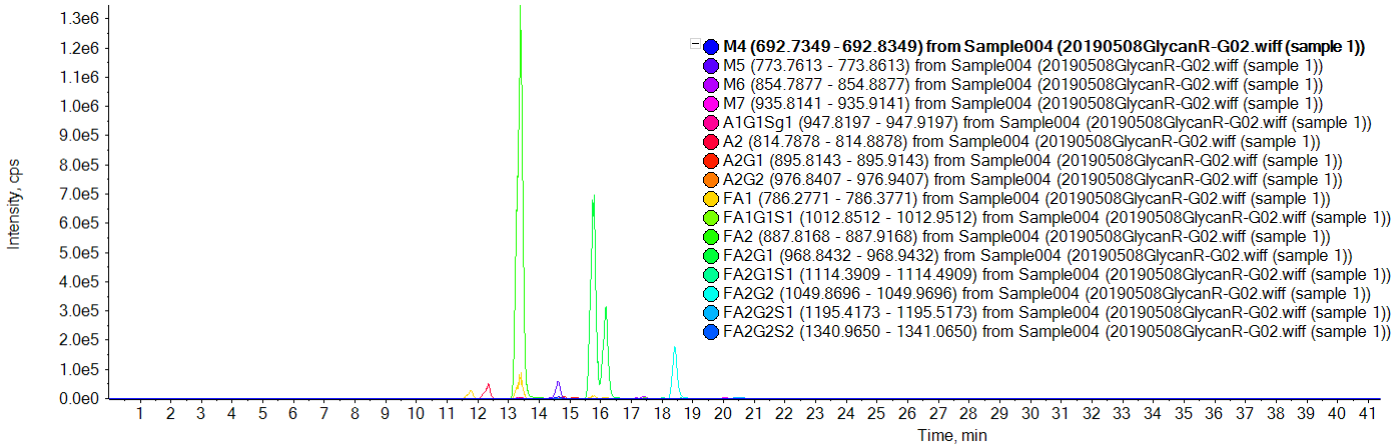

Figure 3. XICs of 16 N-glycans identified in the rituximab mAb sample. XICs were automatically generated according to the *m/z* information from the constructed N-glycan library.

#	Analyte Peak Name	Component Name	Area	Retention Time	Formula	Precursor Mass	Found At Mass	Mass Error (ppm)	Isotope Ratio Difference
2	M4	M4	6.883e+04	14.61	C57H89O32N7	692.785	692.7854	0.8	2.8
3	M5	M5	6.307e+05	14.61	C63H99O37N7	773.811	773.8121	1.0	10.3
4	M6	M6	7.873e+04	17.41	C69H109O42N7	854.838	854.8384	0.8	4.2
5	M7	M7	5.089e+04	20.05	C75H119O47N7	935.864	935.8644	0.3	8.6
13	A1G1Sg1	A1G1Sg1	1.473e+04	13.41	C76H119O46N9	947.870	947.8687	-1.1	12.0
14	A2	A2	5.839e+05	12.34	C67H105O37N9	814.838	814.8371	-1.0	10.5
28	A2G1	A2G1	8.594e+04	14.80	C73H115O42N9	895.864	895.8646	0.4	3.8
32	A2G2	A2G2	1.326e+04	17.53	C79H125O47N9	976.891	976.8900	-0.7	11.5
88	FA1	FA1	9.456e+05	13.38	C65H102O36N8	786.327	786.3265	-0.8	3.0
92	FA1G1S1	FA1G1S1	3.196e+04	17.43	C82H129O49N9	1012.901	1012.9019	0.7	3.8
94	FA2	FA2	1.655e+07	13.38	C73H115O41N9	887.867	887.8674	0.6	10.2
106	FA2G1	FA2G1	8.241e+06	15.76	C79H125O46N9	968.893	968.8940	0.8	13.4
108	FA2G1S1	FA2G1S1	2.919e+04	18.02	C90H142O54N10	1114.441	1114.4405	-0.4	4.6
110	FA2G2	FA2G2	1.937e+06	18.41	C85H135O51N9	1049.920	1049.9206	0.9	12.5
115	FA2G2S1	FA2G2S1	4.096e+04	20.40	C96H152O59N10	1195.467	1195.4677	0.3	2.3
116	FA2G2S2	FA2G2S2	1.470e+04	22.58	C107H169O67N11	1341.015	1341.0139	-0.8	3.6

Figure 4. Report summarizing 16 identified N-glycans in the rituximab mAb sample. The table was generated through the report function of SCIEX OS software.

Conclusions

- The established N-glycan library enables scientists to perform fast, comprehensive and reliable glycan analyses for protein therapeutics.
- Confidence in the workflow is achieved through direct access of all relevant information, such as XIC and retention time, isotopic patterns of the precursors and the related MS/MS spectrum within SCIEX OS software.
- Robust data processing for released glycan analysis—covering XIC extraction, glycan identification, data filtering and report generation—is covered in a single software platform, SCIEX OS software.
- Accurate N-glycan identification and relative quantification is achieved routinely with the data generated by SCIEX accurate mass instrumentation.

References

1. Characterization of 2-AB labelled released N-linked glycans using the X500B QTOF system. SCIEX Technical Note, RUO-MKT-02-9202-A.
2. Characterization of 2-AB labeled released N-linked glycans using SCIEX TripleTOF 5600+ LC-MS/MS system. SCIEX Technical Note, RUO-MKT-02-10500-A.

SCIEX

Headquarters
500 Old Connecticut Path | Framingham, MA 01701 USA
Phone 508-383-7700
sciex.com

International Sales
For our office locations please call the division headquarters or refer to our website at
sciex.com/offices