Characterizing Reproducibility of Glycoform Distributions for SARS-CoV-2 Spike Protein-Derived Glycopeptides Across Recombinant Protein Sources Using Automated, Mass Spectral Library-Based Methods

PRESENTER:
Meghan C. Burke

AIM: To characterize site-specific glycan variability for 11 commercial sources of recombinant SARS-CoV-2 spike protein using glycan abundance distribution spectra (GADS).

METHODS

1. Recombinant proteins were digested using up to seven proteases, or combinations of proteases.
2. LC-MS/MS analysis was performed using a 75 cm Acclaim PepMap RSLC column (Thermo Fisher Scientific) in-line with an Orbitrap Fusion Lumos (Thermo Fisher Scientific).
3. MSMS were acquired using steppedHCD at NCE values of 15, 25 and 35.
4. Ion trap scans were triggered by m/z 204.087.
5. In-house software was used to filter glycopeptide spectrum assignments from Byonic using MS1-level information, such as retention time.

RESULTS

- Distributions for the same site derived from data acquired from different peptide sequences or laboratories illustrate the reproducibility of site-specific glycoform distributions (Figure 1).

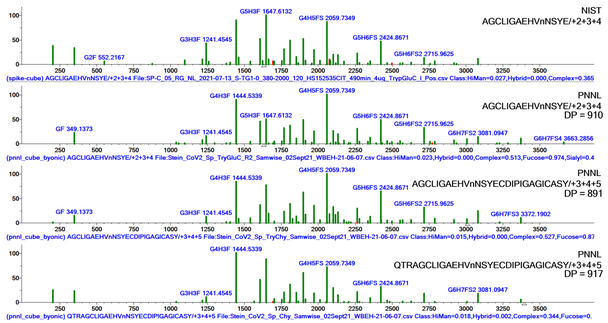
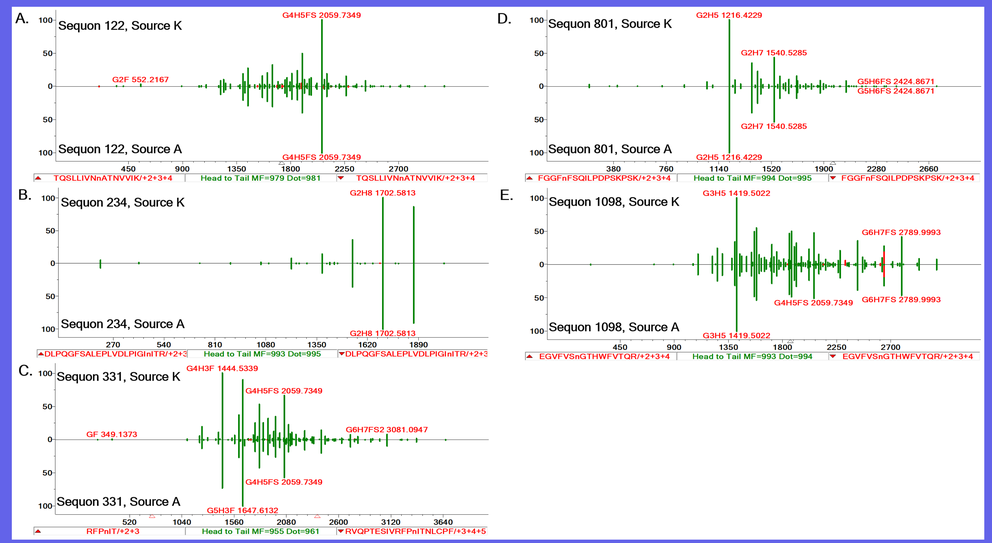


Figure 1. Comparison of GADS derived from Asn657 for data acquired by PNNL and NIST. Dot product (DP) values listed are relative to the GADS for AGCLIGAEHVnNSYE/+2+3+4 from data acquired at NIST.

Recombinant proteins prepared under the ***same conditions*** yield ***reproducible glycoform distributions*** from different commercial vendors.



Source	Sequence	Cells	Furin	Pro Substitution	Mutations	C-term Tag
A	16-1213	HEK293	RAAA			T4, 10xHis
B	16-1213	HEK293	RAAA	F817, A892, A899, A942, K986, K87		T4, 10xHis
C	16-1213	HEK293	RAAA			His
D	1-1273	HEK293 Expi	GSAG	K986, V987		Rho 1D4
E	1-1273	HEK293 Expi	GSAG	K986, V987	del 69-70 & 144, N501Y, A570 D, D614G, P681H, T716I	Rho 1D4
F	15-1208	HEK293	GSAS	K986, K987		6xHis
G	1-1208	CHOExpress	GSAS	F817, A892, A899, A942, K986, V987	del 69-70 & 144-145, N501Y, A570D, D614G, P681H	8xHis
H	1-1208	CHOExpress	GSAS	K986, K987		8xHis
I	15-1208	HEK293	GSAS	K986, K987		His
J	16-1188	HEK293	RAAA			T4, His+Avi
K	16-1213	HEK293	RAAA			T4, 10xHis

Take a picture to **link with** the **full paper and supplemental material**

Table 1. Principal contributors to variability of the most dissimilar replicate GADS. Sources of variation were identified from manual inspection of all 145 inter-replicate GADS with dot products < 850, of which 138 were found attributable to the sources listed.

Description	Count	Percent
Poorly retained glycopeptide	61	44.2
Incorrect monoisotopic assignment	32	23.1
Incorrect ID/ Low abundance/low quality MS2	25	18.1
XIC uncertainty	18	13.0
Probable Ion Suppression	1	0.72
Possible in-source fragmentation	1	0.72

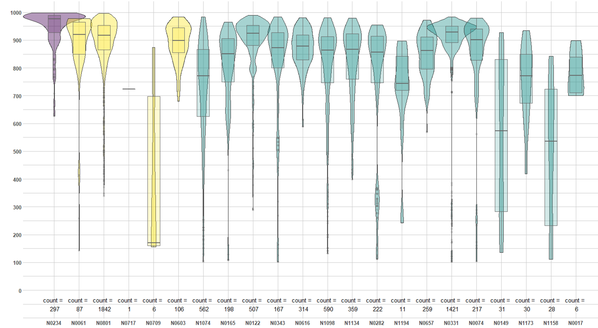


Figure 2. Dot product distribution for GADS corresponding to the same sequon with nSpec ≥ 100 (different sequence/same source). The sequons have been sorted in order of descending total abundance of high mannose glycans. The colors are also used to highlight oligomannose content of 80-100% (purple), 30-79% (yellow) and 0-19% (aquamarine).

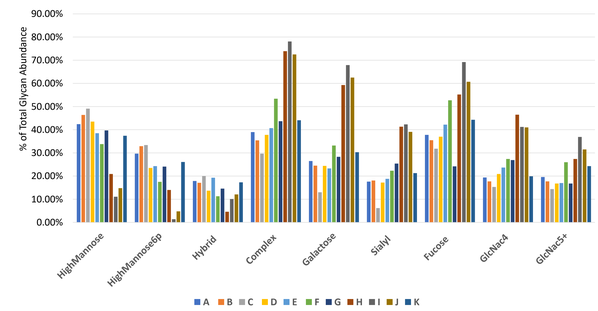


Figure 3. Comparison of glycan class abundance across all 11 sources of recombinant spike protein. Here the classes are defined as high mannose (G(2)H(5-9)), high mannose 6p (G(2)H(6-)), hybrid (G(3)H(5-8)), complex (G(4-6)H(5-7)), galactose-containing (non-high mannose with > 3 hexose), sialylated, fucosylated, and complex-type glycans containing four GlcNAc and ≥ 5 GlcNAc.

Meghan C. Burke, Yi Liu, Concepcion Remoroza, Yuri A. Mirokhin, Sergey L. Sheetlin, Dmitrii V. Tchekhovskoi, Guanghui Wang, Xiaoyu Yang, Stephen E. Stein

Mass Spectrometry Data Center,
National Institute of Standards and Technology,
Gaithersburg, MD

NIST National Institute of Standards and Technology U.S. Department of Commerce | Mass Spectrometry Data Center