The Utility of Nanoparticle Protein Coronas for Studying the Plasma Glycoproteome

Gary M. Wilson[1], Sangtae Kim[2], Shadi Ferdosi[2], and Marshall W. Bern[1]

[1]Protein Metrics Inc., Cupertino, California

[2]Seer Inc., Redwood City, California

PROTEIN METRICS
www.proteinmetrics.com

seer

Introduction

A novel strategy using nanoparticle protein coronas, which enables the detection of low abundance proteins and improved plasma proteome coverage, was recently reported [1]. Nanoparticles enable robust and reproducible measurement of the plasma proteome, which assists in the discovery of novel protein biomarkers. Since nanoparticle coronas differentially interrogate complex samples at the proteoform level (i.e., native proteins) and plasma proteins are often glycosylated, we investigated whether this strategy could also provide a robust route towards the plasma glycoproteome without the need for subsequent enrichment of glycosylated peptides. Protein glycosylation states can provide diagnostic evidence and a complementary insight when the total protein abundance is uninformative. Improved methods for profiling the plasma glycoproteome thus harbor the potential to accelerate biomarker discovery.

Multi Nanoparticle-based Plasma Proteomics

Blood plasma is the ideal biospecimen to assess the health and diseased states of humans since it passes by almost all tissues and is accessible from a large number of individuals at different time points. However, the challenging wide dynamic range of the plasma proteome comprising thousands of proteins and their proteoforms (e.g., PTMs, isoforms) limits unbiased proteomics at depth in large-scale with current technologies. To overcome this limitation, the Proteograph™ Product Suite (Seer Inc.) was developed, which is a fast and scalable technology that employs intricate protein-coronas formed on the surface of engineered nanoparticles (NPs) to interrogate the depth of plasma proteomes. A combination of the selected 5 NP allows rapid quantification of ~ 3,000 proteins across 7 orders of magnitude from a set of plasma with high precision. The key to expand the capability of the NPs in proteomics is to characterize physicochemical properties driving protein corona formation while exploring biological pathways interrogated with each NP.

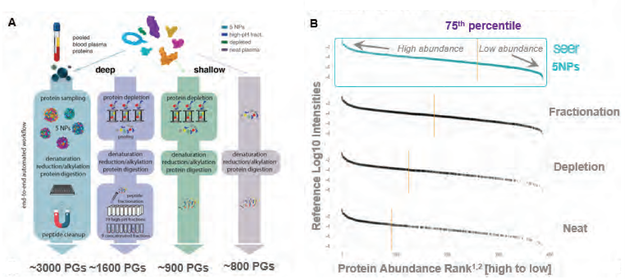
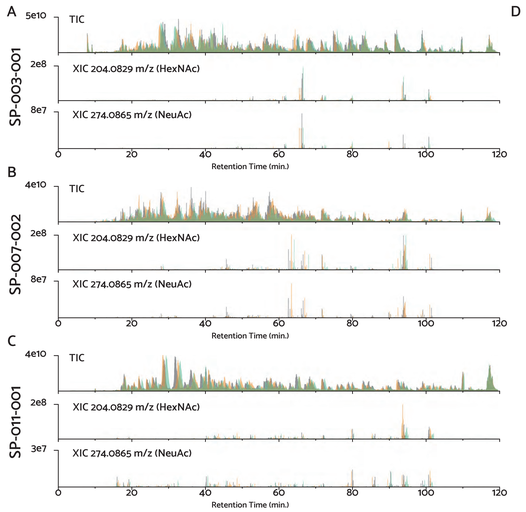
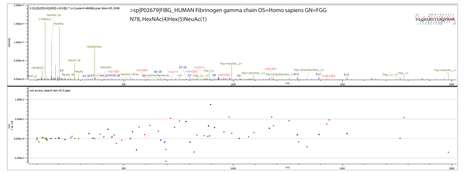


Figure 1. Depth of coverage with different label-free plasma proteomics workflows [4]. A) Conventional label-free plasma proteomics workflows compared to Proteograph Product Suite with a 30 minutes DIA analysis for each 5 NP and total analysis time of 2.5 hrs. B) Proteograph data resulted in ~3000 protein groups identification (1% FDR at protein and peptide level) across 7 orders of magnitude dynamic range with DIA-NN.

Results



D



Oxonium ions, or charged glycan fragments, are often produced during the fragmentation of glycosylated precursors and are especially abundant within spectra that are collected following beam-type collisional-activation (i.e., higher energy collisional dissociation, HCD). We observed that the ions for N-acetylhexosamine (HexNAc) and N-acetylneuraminic acid (NeuAc) are prevalent in these data and often co-occur within the same tandem mass spectra, suggesting that the glycosylated peptides within these data were likely to be sialylated. This notion was confirmed in the results from Byonic searches of the raw data from nanoparticle protein corona enrichment of plasma proteins (Figure 3).

Figure 2: The total ion current (TIC) and the extracted ion current (XIC) for N-acetylhexosamine and N-acetylneuraminic acid oxonium ions from previously published data[1] are shown for A) SP-003-001, B) SP-007-002, and C) SP-011-001. The individual replicates are colored in green, orange, and blue. D) A representative glycopeptide spectrum with high sequence coverage.

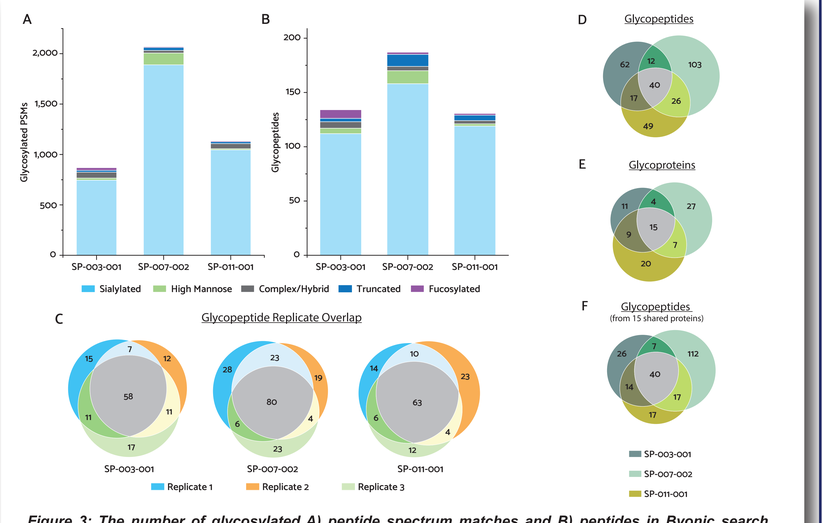
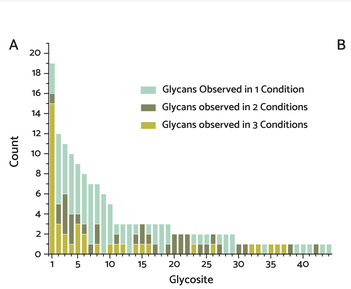


Figure 3: The number of glycosylated A) peptide spectrum matches and B) peptides in Byonic search results are broken down by glycan type. The overlap of glycopeptide identifications between individual replicates is displayed in C) for each nanoparticle. The overlapping identifications between the different nanoparticle data sets is shown for D) glycopeptides, E) glycoproteins, and F) glycopeptides originating from the 15 glycoproteins identified from all three nanoparticle data. We observe that nanoparticle protein coronas provide reproducible and complementary enrichments with >60% overlap within and <30%overlap across each nanoparticle data set.

Protein glycosylation states and protein corona formation



B

ID	Accession	Gene	Glycosite
1	P02679	FIBG	N78
2	P04004	VTNC	N242
3	P04004	VTNC	N86
4	P00734	THRB	N416
5	P02675	FIBB	N394
6	P04004	VTNC	N169
7	P27169	PON1	N253
8	P01042	KNG1	N294
9	P02745	C1QA	N146
10	P01859	IGJ	N176
11	P01591	IGHG2	N71
12	P0C0L5	C04B	N226
13	P12694	ODBA	N308
14	P01042	KNG1	N169
15	P00738	HPT	N184
16	P02765	FETUA	N176
17	P03952	KLKB1	N494
18	P00734	THRB	N121
19	P01042	KNG1	N48
20	P80108	PHLD	N321
21	P03951	FA11	N450
22	Q8WV28	BLNK	N266
23	P10909	CLUS	N374
24	P49908	SEPP1	N119
25	P03952	KLKB1	N453
26	P01009	A1AT	N271
27	P03952	KLKB1	N308
28	P10909	FA5	N354
29	P03952	CLUS	N396
30	P03951	KLKB1	N126
31	P03952	FA11	N127
32	P12259	KLKB1	N1703
33	P49908	FA5	N83
34	P40197	SEPP1	N312
35	P00738	CIR	N241
36	P12259	GPV	N821
37	P00736	HPT	N514
38	P16671	FA5	N417
39	P00734	CD36	N143
40	P12259	FA5	N297
41	P12259	THRB	N1559
42	P0C0L5	C04B	N1328
43	P12259	FA5	N460
44	P10909	CLUS	N145

Figure 4: A) The frequency that the same glycan/glycosite pair is identified across nanoparticle coronas. B) The proteins corresponding to the X-axis in A) are provided as a table.

To assess the influence of glycosylation on protein corona formation, we asked whether the same glycan modifications are observed on the glycosylated sites of the proteins that are commonly observed across different nanoparticles. We observe that only one glycosite displays a conserved microheterogeneity of glycosylation across these nanoparticles: of the 19 glycan modifications identified at N78 of fibrinogen (Uniprot accession: P02679), 15 glycans are consistently identified across these three nanoparticles. In contrast, fewer than 1/3 of the glycans observed in other highly occupied sites are consistently observed across all three nanoparticles, further suggesting that nanoparticle protein coronas can enrich for different proteoform variants. We suspect that the conserved microheterogeneity that we observe is due to its high concentration of fibrinogen in blood plasma (>200 mg/dl).

Conclusions

Together, these data provide evidence that nanoparticle protein coronas provide the ability to analyze subpopulations of the glycoproteome without the need for subsequent, glycopeptide-specific enrichment. As well, the different nanoparticles offer complementary views of the plasma glycoproteome due to their specificities for different proteins, and likely, different glycosylated proteoforms. Further study with fragmentation methods that provide greater coverage of glycosylated precursors will undoubtedly uncover the true extent of the glycoproteomes enriched by these methods and their potential for use in biomarker discovery programs.

References

1) Blume, J.E., Manning, W.C., Troiano, G. et al. Rapid, deep and precise profiling of the plasma proteome with multi-nanoparticle protein corona. Nat Commun 11, 3662 (2020).

2) Keshishian H, Burgess MW, Gillette MA, Mertins P, Clauser KR, Mani DR, Kuhn EW, Farrell LA, Gerszten RE, Carr SA. Multiplexed, Quantitative Workflow for Sensitive Biomarker Discovery in Plasma Yields Novel Candidates for Early Myocardial Injury. Mol Cell Proteomics. 2015 Sep;14(9):2375-93.

3) Keshishian, H., Burgess, M., Specht, H. et al. Quantitative, multiplexed workflow for deep analysis of human blood plasma and biomarker discovery by mass spectrometry. Nat Protoc 12, 1683–1701 (2017).

4) Ferdosi, S., Tangeysh, B., Brown, T. et al. Engineered Nanoparticles Enable Deep Proteomics Studies at Scale by Leveraging Tunable Nano-bio Interactions, *under review*.

This work was sponsored in part by NIH grant 1 R41 GM142363-01.