The Utility of Nanoparticle Protein Coronas for Studying the Plasma Glycoproteome

Gary M. Wilson1, Sangtae Kim2, Shadi Ferdosi2, and Marshall W. Bern1
1Protein Metrics Inc., Cupertino, California
2Seer Inc., Redwood City, California

Introduction
A novel strategy using nanoparticle protein coronas, which enables the detection of low abundance proteins and improved plasma proteome coverage, was recently reported. 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 critical 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 through almost all tissues and is accessible from a fraction 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.

Results
To assess the fluorescence 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 glycoprotein displays microheterogeneity of glycosylation across these nanoparticles: 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

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

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 20 minutes 2Dx analysis for each 3 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-MSX.

Figure 2: The total ion current (TIC) and the extracted ion current (XIC) for N-acetylhexosamine and N-acetylneuraminic acid (NeuAc) are shown for A) SP-011-001, B) SP-017-002, and C) SP-017-003. 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 in 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.

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.

Figure 5: A) A representative glycopeptide spectrum with high sequence coverage from the Protein Metrics database. B) Byonic searches of the raw data from nanoparticle protein coronas. C) The proteins corresponding to the X-axis in A) are provided as a table.