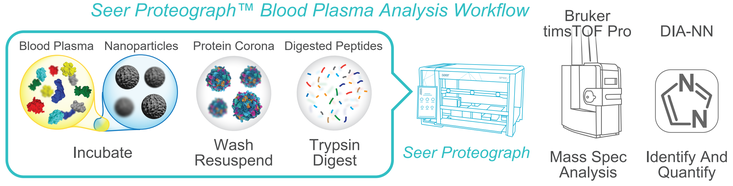
Dissecting the dynamics of protein corona formation on nanoparticles allows reconstructing deep plasma protein concentrations and discovering novel proteoforms

Alexey Stukalov*, *Shadi Ferdosi, Yingxiang Huang, Moaraj Hasan, Brittany Lee, Serafim Batzoglou, Asim Siddiqui, and Daniel Hornburg*

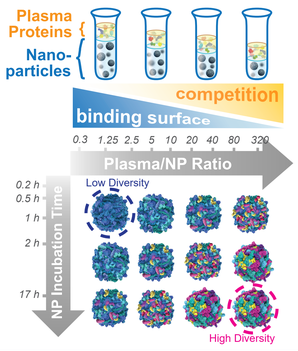
Blood plasma is a rich and accessible source of known and potential novel **biomarkers**. However, the range of protein concentrations in plasma spans **10 orders of magnitude**, and physiological roles of proteins in the low abundance fraction are largely unexplored. Even less is known about the **proteoform** composition of these proteins.

Proteograph™ Product Suite with Seer's proprietary **nanoparticles (NP)** enable deep, unbiased and scalable plasma analysis[1,2,3]. **Protein corona** that forms around a nanoparticle has compressed dynamic range, which brings low abundance proteins within detection limits of MS instruments. By profiling the dynamics of corona formation, we revealed a wide diversity of **NP–protein interaction** modes[1]. Here, we use these data to reconstruct neat **concentrations** of proteins, while peptide-level analysis identifies known and novel proteoforms.

Seer Proteograph™ Blood Plasma Analysis Workflow



Profiling the Formation of Nanoparticle Protein Corona



To understand the dynamics of NP–protein interactions, we profiled protein coronas across a wide range of conditions[1]:

- ***3*** different Seer nanoparticles
- ***9*** plasma/nanoparticle ratios
- ***5*** nanoparticle incubation times

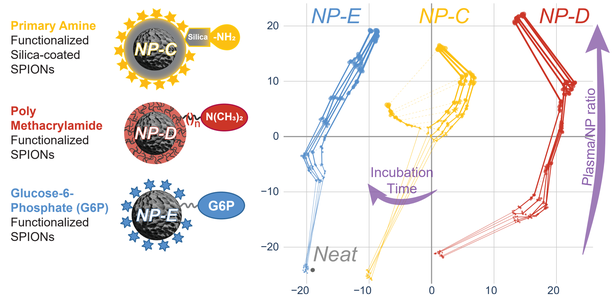
By modulating conditions, we shift the dynamic equilibrium of NP–proteins system, and the composition of protein corona changes.

In total we have measured ***534*** protein corona samples by LC-MS using *Bruker timsTOF Pro* in *diaPASEF* mode. The MS data were analyzed with *DIA-NN* 1.8 yielding intensity profiles for ***3184*** protein groups.

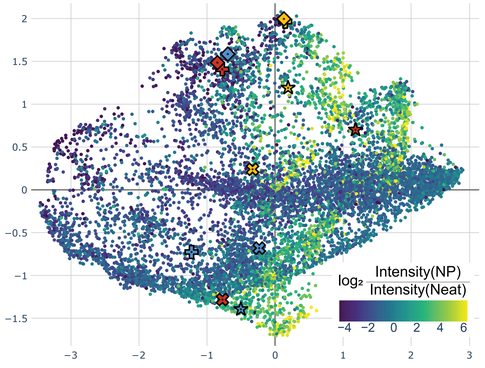
The analysis of these data confirmed high reproducibility and consistency of the Seer NP workflow.

Diversity of Nanoparticle-Protein Interactions

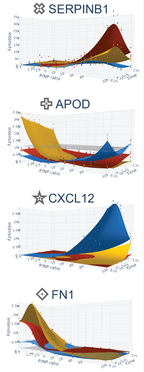
UMAP of Protein Corona Experiments



UMAP of NP-Protein Profiles

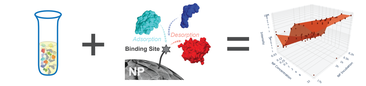


NP-Protein Profile Examples



Reconstructing Protein Abundance Using NP-Protein Interaction Profiles

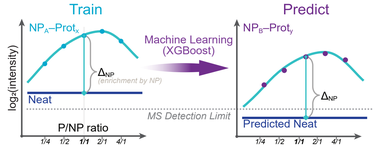
Protein changes in NP coronas accurately follow protein concentration changes between the biosamples, but the dynamic range compression affects absolute abundance estimates.



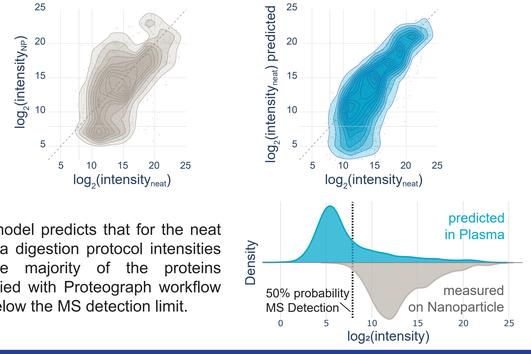
With machine learning applied to the NP-Protein profilles we can reconstruct neat plasma concentrations.

Using ML to Predict Protein Enrichment in NP

Similar NP-Protein profiles indicate similar Proteograph Assay vs neat digestion workflow enrichment.



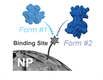
ML Enhances Absolute Neat Protein Intensity Estimates



The model predicts that for the neat plasma digestion protocol intensities of the majority of the proteins identified with Proteograph workflow are below the MS detection limit.

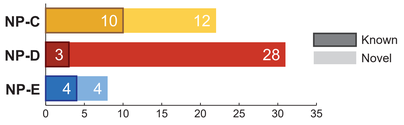
Detecting Proteoforms In NP-Peptide Profiles

Alternative splicing, posttranslational cleavage and other modifications are the reasons the same protein can exist in multiple forms – *proteoforms*. These forms may have distinct biological roles and differ in biochemical properties that manifest themselves in different nanoparticle corona intensity profiles of their peptides.



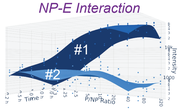
Initial correlation analysis of the peptide-level profiles using the COPF method[4,5] has identified 61 protein groups composed of several proteoforms, 44 of those groups contained proteoforms not described before.

Detected Proteoforms

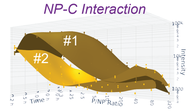


Examples

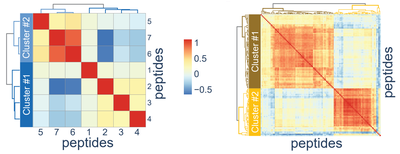
Progranulin



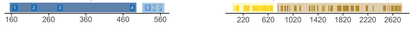
von Willebrand Factor



Pearson Correlation of log_2 Intensities



Peptide Positions on Protein Sequence



Conclusions

- Understanding the NP–protein interactions on the quantitative level opens new possibilities for the nanoparticle-enabled LC-MS/MS analysis.
- NP protein corona profiling allows to streamline the design of nanoparticles interrogating specific physicochemical fraction of the proteome occupied by hundreds to thousands of proteoforms.
- NP-protein profiling improves the processing of the retrieved data, including large cohort studies.

References

1. Ferdosi, Stukalov et al, **Advanced Materials** (2022)
2. Blume et al, **Nat Comm** (2020)
3. Ferdosi et al, **PNAS** (2022)
4. Donovan et al, **PLoS ONE** (2023)
5. Bludau et al, **Nat Comm** (2021)

Publications