NAUTILUS™

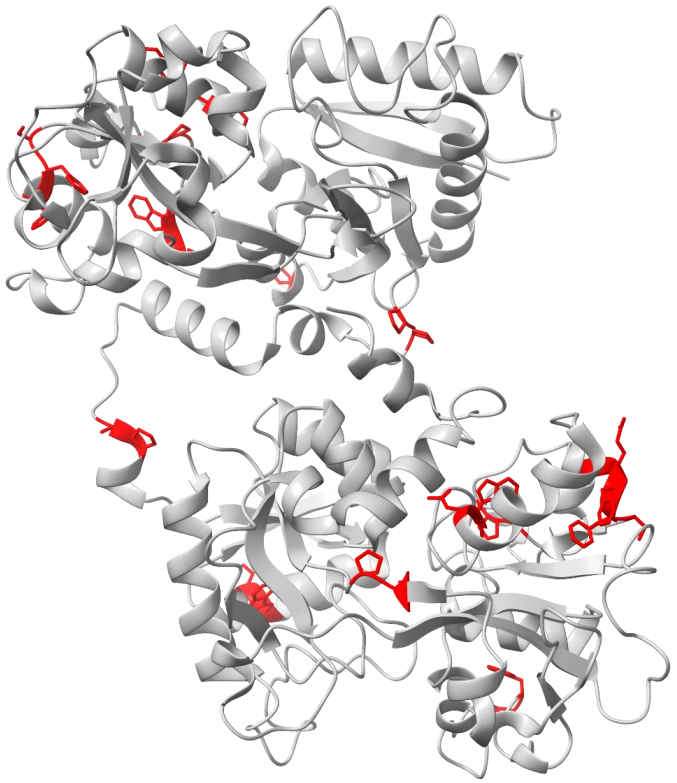
BIOTECHNOLOGY

Experimental Findings Presented at US HUPO Annual Conference

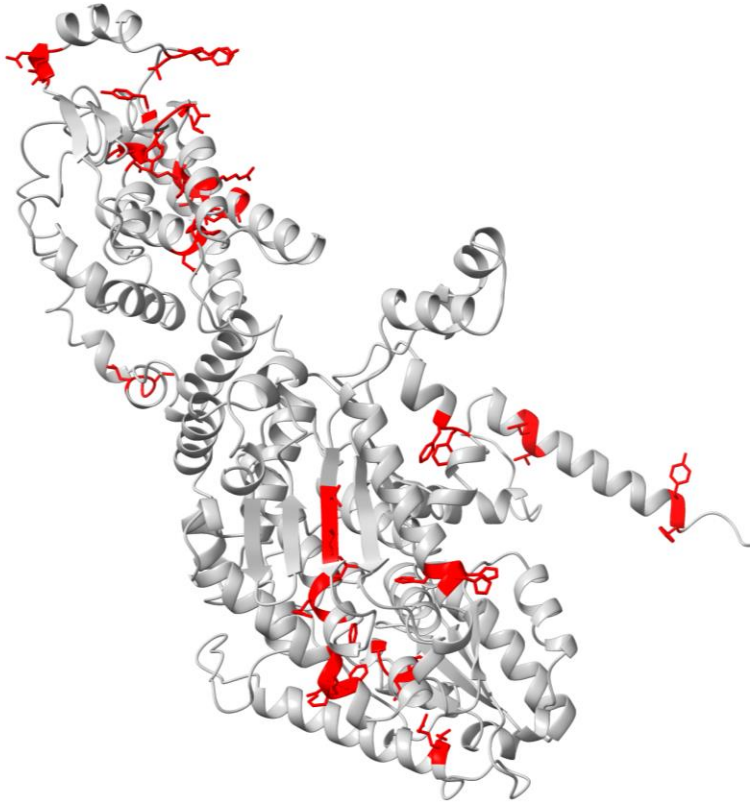
MARCH 11, 2024

Multi-affinity probes bind to buried/structured regions in proteins

- Binding of a multi-affinity probe to its intended peptide target is required
- Multi-affinity probes must also be able to bind to short epitopes in full-length proteins
- We have demonstrated binding to short epitopes in internal portions of full-length proteins (red) using both aptamer- and antibody-based multi-affinity probes



Transferrin



G6PI

NAUTILUS™ BIOTECHNOLOGY

False discovery rate in PrISM

Identification	P-val
KRAS	.99998
P53	.9995
⋮	⋮
SMP1	.998
DECOY	DECOY
EGFR	.998

We apply a target-decoy based method to estimate what % of identifications are false

- Generate a database with decoy proteins that *look* like real proteins
- Analyze how often decoy IDs occur to estimate the false ID rate

In our system, the most likely failure mode occurs when proteins with high sequence similarity are mistaken as one another

Therefore, we want to generate decoys that:

- Reflect the sequence structure of real proteins
- Capture the relative likelihood of one protein being mis-identified as another

We demonstrated that decoys either generated with shuffled protein sequences or alternative proteomes effectively estimated the false discovery rates

We also demonstrate that the FDR estimation is not impacted by possible failure modes of the system, including mis-prediction of the binding rates between probes and proteins from the proteome

NAUTILUS™ BIOTECHNOLOGY

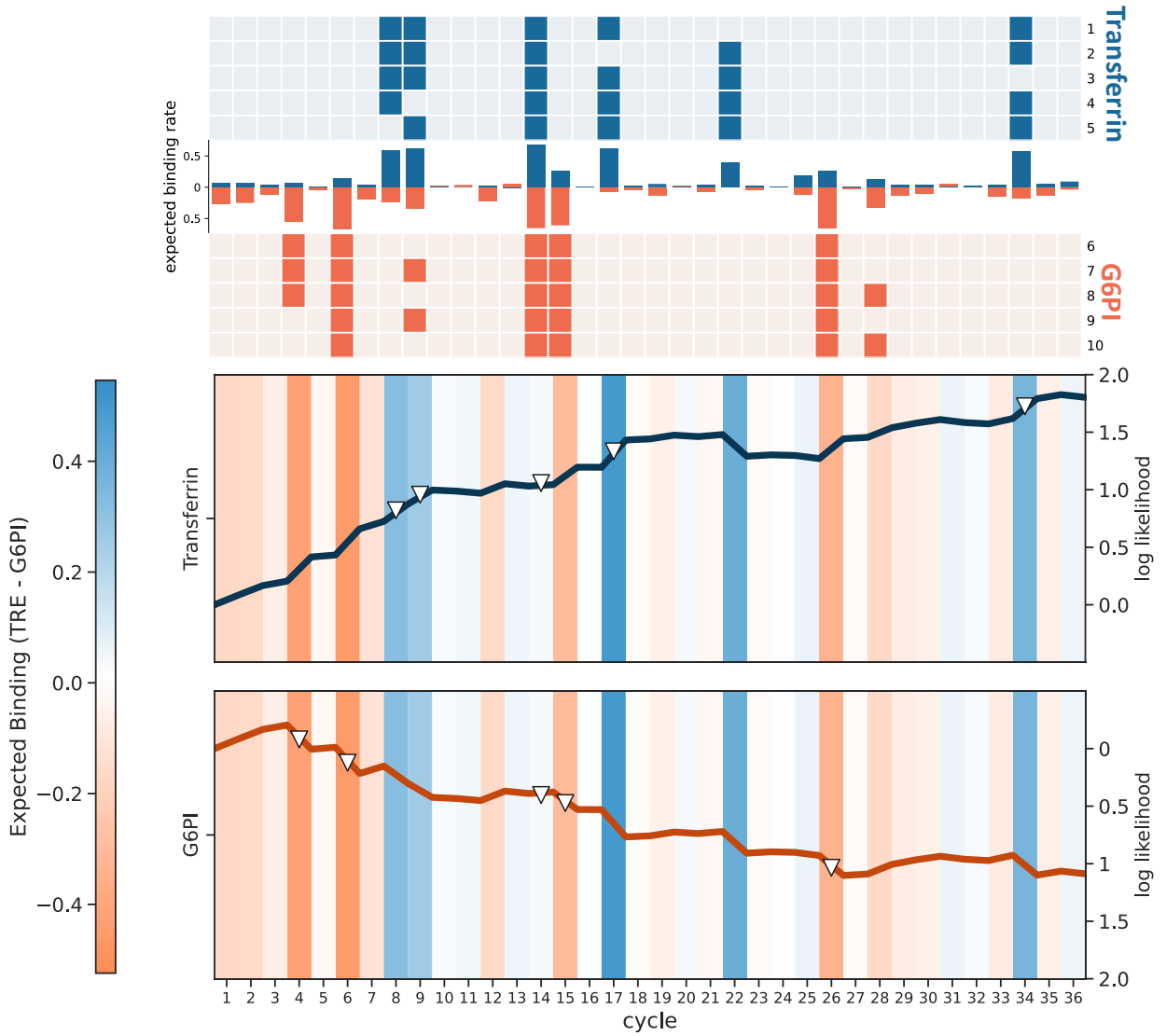
Binding patterns to transferrin and G6PI at single-molecule resolution

Samples consisting of transferrin (TRFE), glucose 6 phosphate (G6PI), pyruvate kinase M2 (PKM2), a model protein, no protein (negative control) or mixtures thereof were deposited into flowcell lanes for PrISM analysis.

Shown top are the 5 most prevalent binding patterns from these experiments for transferrin and G6PI.

From these binding patterns, machine learning tools identify each molecule. Each additional cycle builds additional information about protein identity as transferrin and G6PI have different binding patterns, indicated by the triangles.

The resulting difference in probability between the best-matching protein, and the next best protein in the database leads to confident protein identifications.



Ultra-sensitive quantification of transferrin

(TOP-LEFT) Sensitive single-molecule measurements: Lower limit of detection in the high-yocto to low-zeptomole range

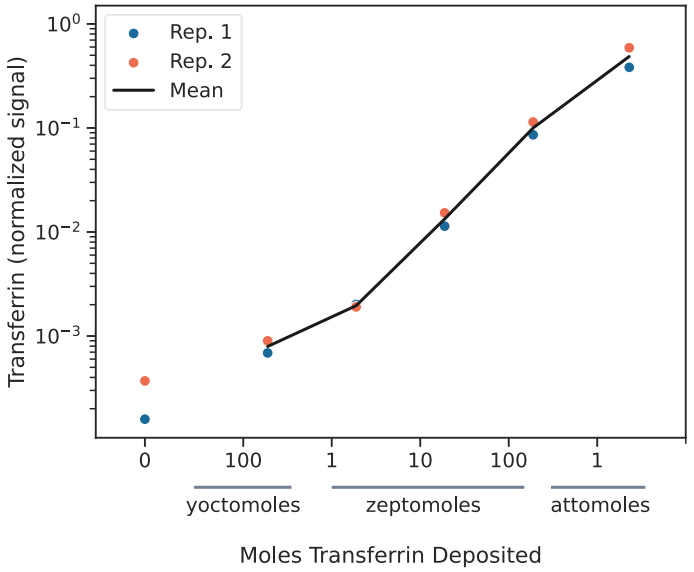
Reproducible single-molecule measurements: The highest abundance Transferrin measurement (2 attomoles) was repeated 7 times across 4 days and 7 flow cells with a CV of **7.7%**

(TOP-RIGHT) High fidelity quantification: Multi-cycle decoding data is significantly more sensitive and error tolerant than achievable with any one multi-affinity probe alone

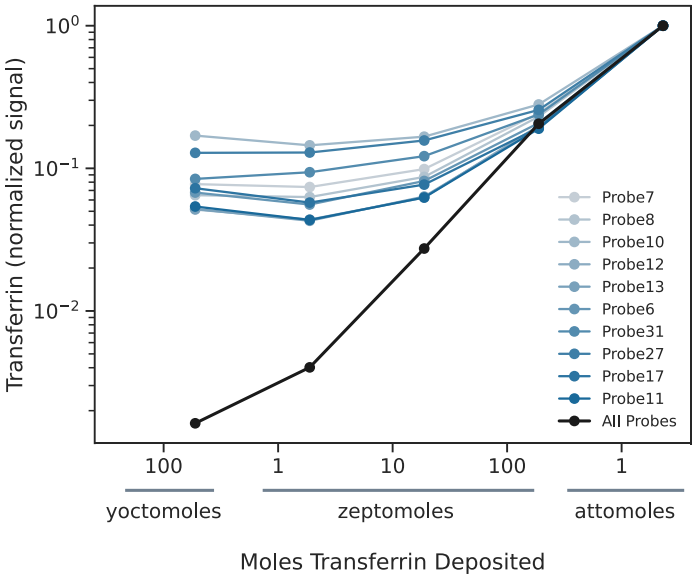
This improvement arises from the ability of the machine learning software to better identify proteins whose identifications were derived from either false positive or false negative bindings

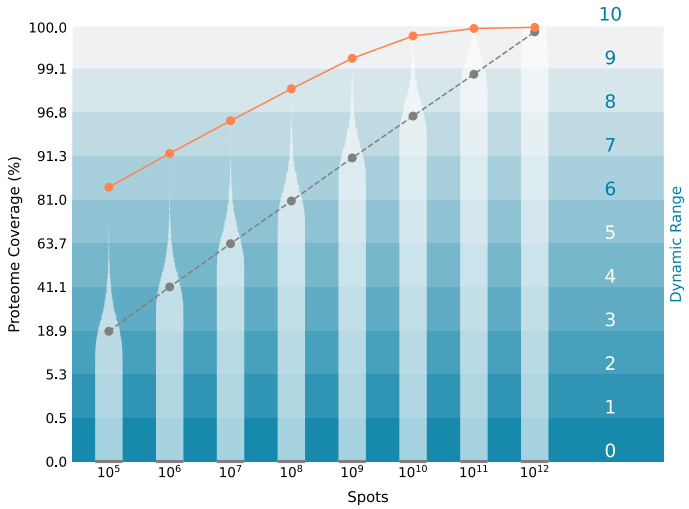
(BOT-LEFT) : As one increases the number of molecules measured with larger flowcells, one increases the dynamic range of the platform

Transferrin dilution series in a background of alternate protein, or null scaffolds



Multi-probe decode vs single-probe

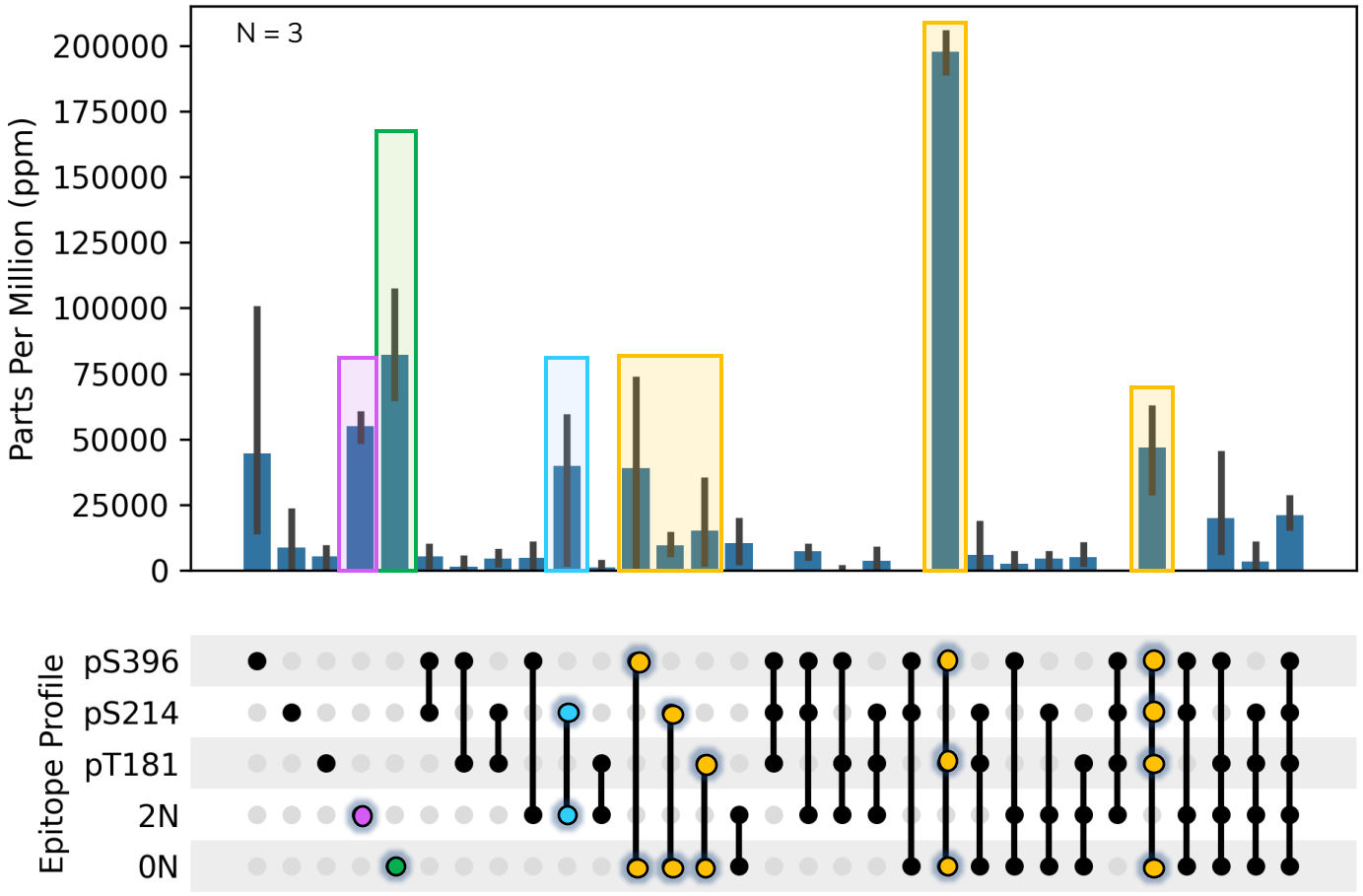




In orange is shown the dynamic range as the difference between the lowest abundance and highest abundance protein.

In grey is shown the dynamic range where greater than 90% of the proteins at a given concentration are measured.

Quantification of mixtures of proteoforms



Tau proteoforms	Molar ratio
0N	25
0N ERK (181 & 396)	50
2N	12.5
2N PKA (214)	12.5

Exploiting the massively parallel nature of our platform, the relative abundances of seven Tau proteoforms were accurately quantified. This measurement is intractable on both traditional and emerging peptide-based platforms.

We additionally showed how the platform can be applied to measure EGFR proteoforms.

NAUTILUS™ BIOTECHNOLOGY

NAUTILUS™

BIOTECHNOLOGY

Corporate Headquarters
2701 Eastlake Avenue, Suite 300
Seattle, WA 98102

Research Headquarters
835 Industrial Rd, Suite 200
San Carlos, CA 94070

(206) 333-2001 info@nautilus.bio www.nautilus.bio