Protein Sequencing with Single Amino Acid Resolution Discerns Peptides that Discriminate Tropomyosin Proteoforms

Natchanon Sittipongpittaya[1*]; Madison M. Mehlferber[1], Kenneth A. Skinner[2];* Erin D. Jeffery[1]; Emily F. Watts-Whitehead[1]; Gloria Sheynkman[1]

[1]University of Virginia, Charlottesville, VA
[2]Quantum-Si Incorporated, Branford, CT

*Equal contribution

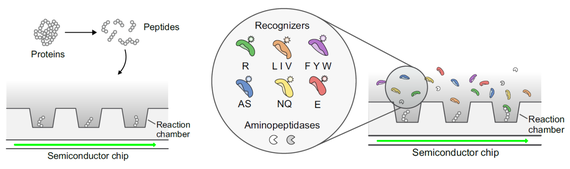
Poster # P01.25

NAA recognizers and aminopeptidases enable parallel sequencing of single peptides

Platinum® instrument

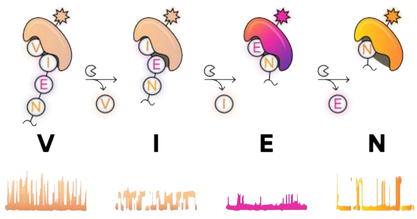
Dimensions:
19.45 x 8.46 x 9.91 in
Weight: 27 lbs

Prepare samples for sequencing



- Peptide functionalization at C-terminal lysines enables immobilization to the chip.
- Labeled N-terminal amino acid (NAA) recognizers and aminopeptidases sequentially read and cleave each AA.
- Sequencing process proceeds in real time without fluidic exchange of reagents.

Sequencing on Platinum

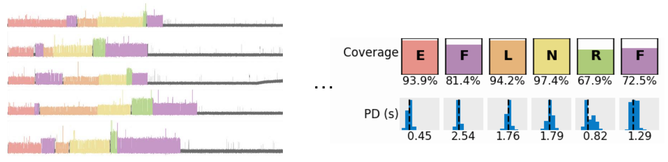


Recognition segments (RSs)

Recognition events:
Pulse duration (PD) can distinguish NAAs with the same recognizer

Aminopeptidase activity:
Cleavage events stochastic at single-trace level

Cloud-based analysis provides kinetic signatures of single AAs



Kinetic signatures

- Measurable characteristics of the series of dynamic recognizer events that uniquely identify a peptide.
- Sensitive to AA variations and PTMs.

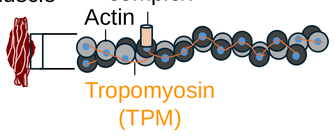
- **Pulse duration (PD)/ Interpulse duration (IPD)** reflects the differences in binding affinity driven by different *dissociation/association* rates for each type of recognizer-NAA interaction.
- Histograms represent the statistical distribution of kinetic data for all the pulses associated with a specific residue. Values reported are the median of the mean PD/IPD for each RS.

TPM variants are highly homologous and functionally distinct proteoforms

>40 TPM isoforms in mammals
Need to detect at peptide/protein level

Skeletal muscle

Troponin complex

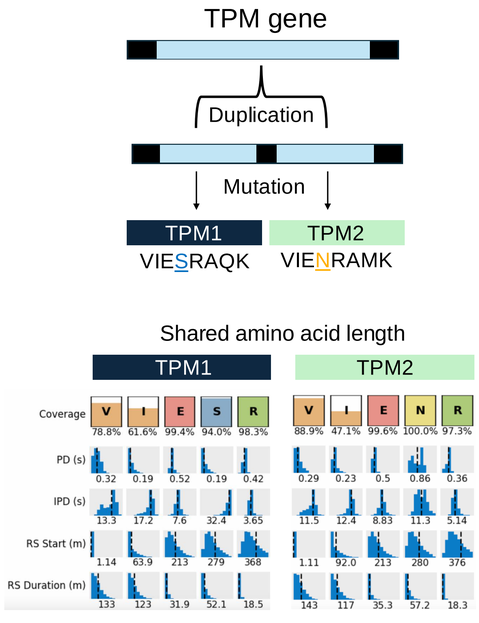


TPM1 and TPM2 share 87% AA sequence identity
Detecting unique peptides can be used to discern proteoforms

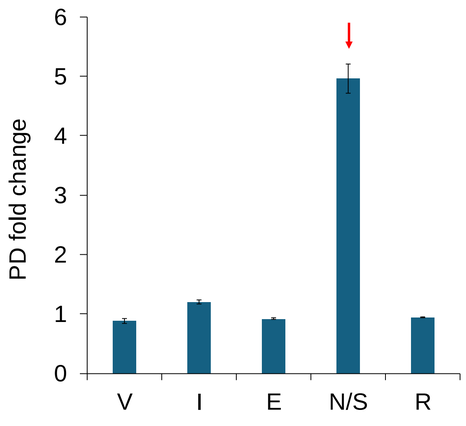
TPM1 paralog
…VIESRAQK…

TPM2 paralog
…VIENRAMK… Region interacts with troponin
…VIENRAMK… Arg mutation in myopathies

Platinum sequences paralog-specific peptides

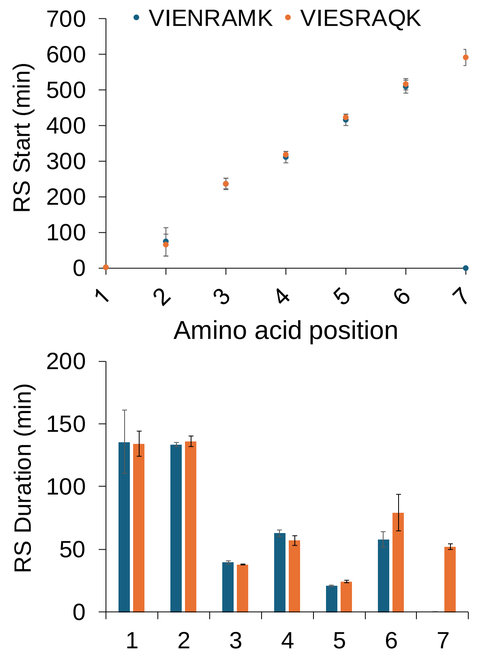


Differential PDs at single position

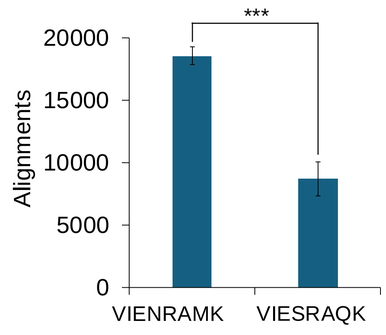


- 5X higher PD with N/Q recognizer vs. S/A recognizer in analogous position
- Higher PD reflects tighter binding and slower dissociation rate

Consistent RS and aminopeptidase activity

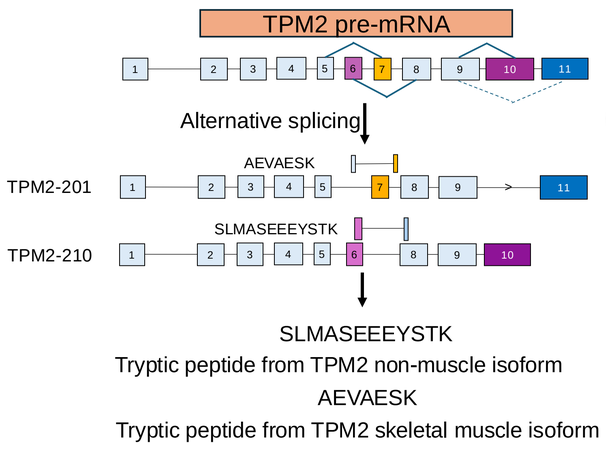


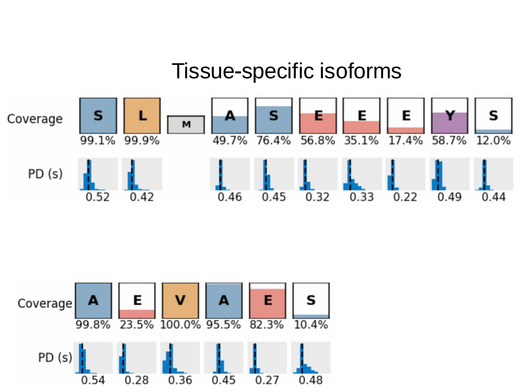
Toward "Quantum-typicity"

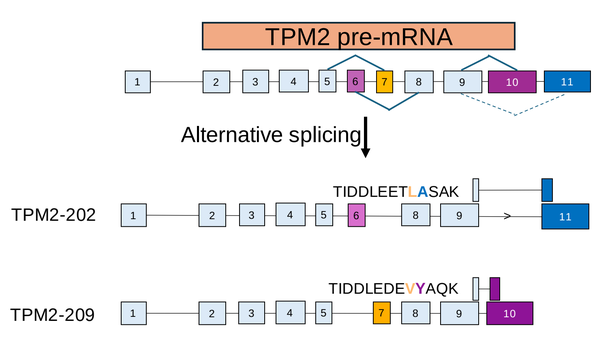


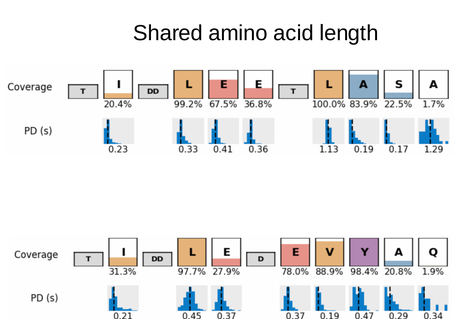
- In proteomic analyses, there is a need to annotate peptides with detectability and isoform specificity for protein ID in mixtures
- Detection of "Quantum-typic" peptides can be used for targeted analysis and relative quantitation of peptides on Platinum

Platinum sequences spliceoform (isoform)-specific TPM2 peptides



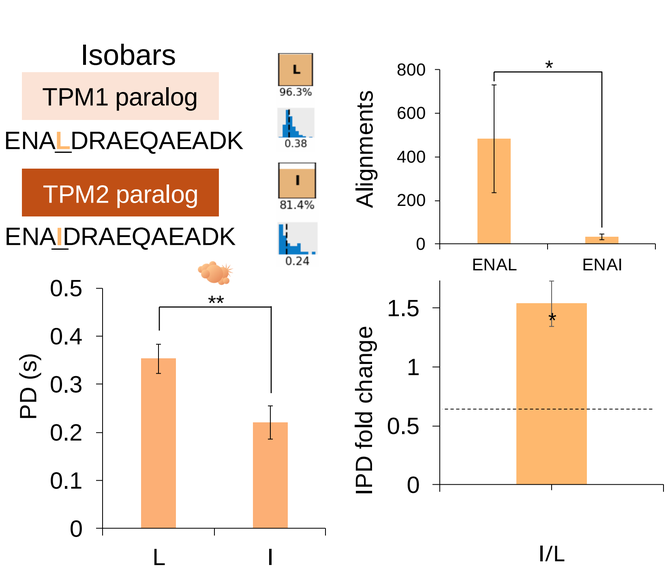






Application: Detect tissue-specific spliceoform expression

Platinum distinguishes isobaric AA variants and unmodified and phosphorylated tyrosine

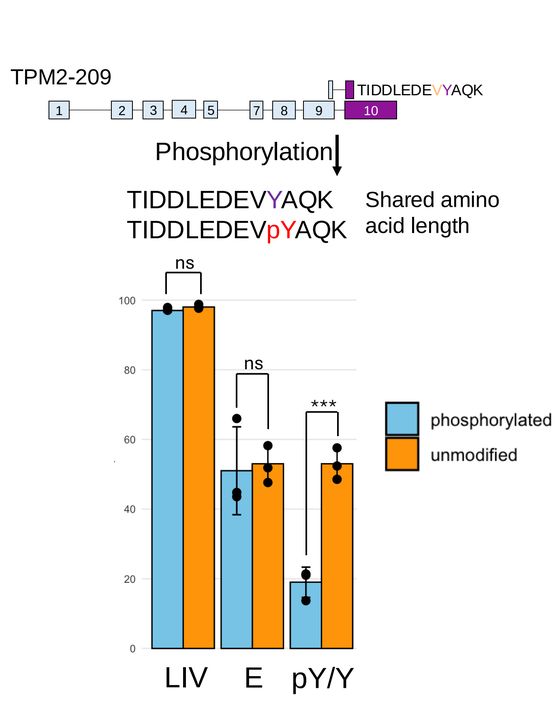


LIV recognizer:

Higher PD and lower IPD with ENAL

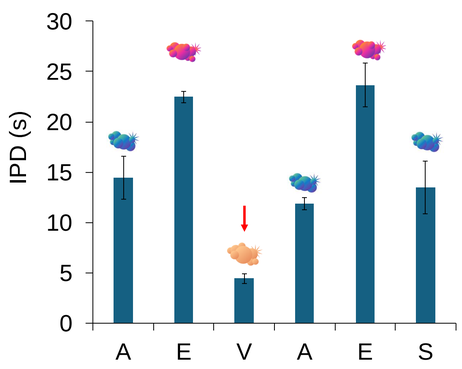
Antibody	Peptide	Isoforms
TM311	-ENALDRAEQAEADK-	TPM1-3

Application: Epitope mapping

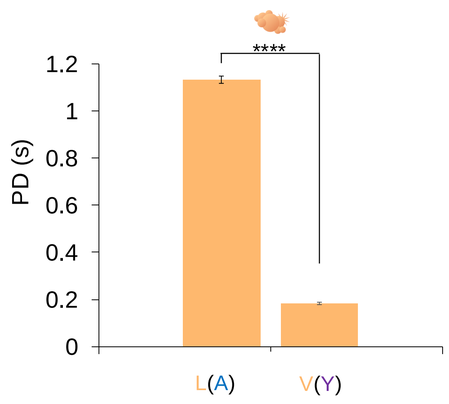


Application: Kinase substrate profiling

NAA recognizers distinguish similar AAs via binding kinetics



- A shorter IPD indicates faster recognizer-NAA association.
- For AEVAESK peptides, IPD profiles are consistent for the recognizers A/S and E, but distinct for V.
- Thus, a single AA change (V) distinguishes AEV and SLM peptides



- A longer PD indicates longer recognizer-NAA residence time
- Platinum can discern L and V with the same recognizer based on kinetics

Summary and Conclusions

- Next-Generation Protein Sequencing (NGPS) resolves proteotypic peptides representing the three main types of regulation — genetic, splicing, and post-translational modifications — contributing to proteomic molecular diversity
- Kinetic signatures produced by NGPS are a novel orthogonal datatype that can complement other commonly used proteomics approaches
- NGPS data can be used for a range of protein characterization applications, including detecting tissue-specific variant expression, epitope mapping, and PTM profiling

References

Reed, Brian D., *et al.* "Real-time dynamic single-molecule protein sequencing on an integrated semiconductor device." *Science* 378.6616 (2022): 186-192.

Schevzov, G., Vrhovski, B., Bryce, N.S., Elmir, S., Qiu, M.R., O'neill, G.M., Yang, N., Verrills, N.M., Kavallaris, M. and Gunning, P.W., 2005. Tissue-specific tropomyosin isoform composition. *Journal of Histochemistry & Cytochemistry*, *53*(5), pp.557-570.

Abood, A., Mesner, L.D., Jeffery, E.D., Murali, M., Lehe, M.D., Saquing, J., Farber, C.R. and Sheynkman, G.M., 2024. Long-read proteogenomics to connect disease-associated sQTLs to the protein isoform effectors of disease. *The American Journal of Human Genetics*, *111*(9), pp.1914-1931.

Disclosures

Kenneth Skinner is a shareholder and employee at Quantum-Si.

Acknowledgments

We thank Bonnie Lun and Meredith Carpenter (Quantum-Si) for technical assistance.

See preprint on biorxiv for more!
https://www.biorxiv.org/content/10.1101/2024.11.04.621980v1

Trademarks/Licensing