Docket #: S22-476

Single-molecule Peptide Sequencing Enabled by Intramolecular DNA Encoded Edman Degradation

Stanford researchers have developed a next-generation protein sequencing platform capable of identifying all the proteins in a cell at single amino acid resolution.

Methods for single-cell RNA and DNA sequencing have revolutionized science and medicine. However, genetic data poorly reports on the proteins in a cell, as post-transcriptional modifications critical to cellular function (e.g., protein phosphorylation, glycosylation, ubiquitination) are not encoded in mRNA, and mRNA levels are often poorly predictive of protein levels. Despite this, robust methods for identifying and quantifying the proteins in a cell are lacking. Current mass spectrometry approaches can only identify a small fraction of all the proteins in a cell, while emerging approaches such as nanopore and real-time dynamic protein sequencing suffer from poor generalizability, throughput, and sensitivity.

Researchers therefore developed a new protein sequencing method that combines the well-established technologies of Edman degradation and high-throughput DNA sequencing. Briefly, one amino acid at a time is removed from the end of each protein and barcoded with DNA that are specific to the protein molecule. The DNA barcoded amino acid derivatives are recognized by an antibody specific for that amino acid. Importantly, such antibodies can also recognize specific post-translational modifications. DNA barcodes associated with each antibody encode the identity of the amino acid, the protein it came from, and its location in the protein. DNA sequencing of these barcodes then reveals the amino acid sequence of each protein. Critically, signal amplification during DNA sequencing allows for high sensitivity.

Stage of Development
Proof of concept: development of Edman degradation conditions compatible with

DNA and DNA-barcoded antibodies specific for certain amino acids (including post-transcriptionally modified amino acids)

Applications

- Single-cell de novo protein sequencing at single-molecule resolution
- High throughput database-assisted single-cell protein identification
- Basic proteomics research
- Diagnostics (e.g., detection of aberrant proteins such as beta-amyloid variants in Alzheimer's disease)
- Target discovery
- Drug development (e.g., targeted protein degradation, de-ubiquitination, phosphorylation, etc)

Advantages

- Combines the well-established methods of Edman degradation and high throughput sequencing
- Potential to sequence all the proteins in a single cell at single molecule resolution
- High sensitivity
- High throughput
- Can identify any arbitrary protein sequence
- Can detect post-translational modifications (e.g., phosphorylation, glycosylation, methylation, acetylation, etc.)

Innovators

- H. Tom Soh
- Liwei Zheng
- Yujia Sun
- Brandon Wilson
- Joseph Buonomo

Licensing Contact

Eileen Lee

Email