



# A NEW ERA OF PROTEIN SEQUENCING

**Breakthrough Next-Generation Protein Sequencing™ technology revolutionizes proteomics, enabling all new insights into protein structure and function**

**T**he study of proteins is crucial to understanding numerous biological processes, and the sequencing of proteins plays a foundational role in this understanding. A protein's unique amino acid sequence—including variants and modifications—dictates structure and function. As research trends toward multiomics and systems-based approaches, greater emphasis has been placed on identifying the actual proteins that are expressed, their proteoforms, and abundance levels rather than the theoretical presence determined through DNA and RNA sequences. This is important due to the differing protein lifespans

and translational efficiencies in addition to the many interactions, movements, and modifications that occur post-translation.

While numerous approaches exist for verifying the presence or absence of an expected protein, confidently identifying the protein responsible for a specific biochemical activity and detecting an unknown protein remains a lengthy and complex task. Yet scientists rely on this knowledge to discover novel biomarkers, better comprehend intricate cellular pathways and disease mechanisms, and create more effective therapies.

While the field of genomics has benefited from a



period of rapid advancement and discovery, protein sequencing technology has trailed behind due to the inherent complexity and diversity of proteins. The 20 amino acids that form protein sequences add considerable technical complexity to sequencing compared to the four nucleotide bases in DNA. Further, DNA sequencing techniques like PCR amplification and sequencing by synthesis depend on the reverse complement of double-stranded DNA, a process not applicable to proteins. Instead, sequences are often deduced from genomic or transcriptomic data, which may not reflect the protein's true sequence due to their naturally low abundance and propensity for modifications. Consequently, there's a need for highly sensitive sequencing technology capable

of directly detecting amino acid sequences and modifications, without relying on amplification or synthesis.

## TRADITIONAL PROTEIN ANALYSIS

Existing methodologies for protein analysis include mass spectrometry (MS) and affinity-based assays, which typically rely on indirect methods for detecting the protein sequence, in addition to sequencing methods like Edman degradation. MS measures mass-to-charge ratios to detect, identify, and quantify peptides. This technique can present identification challenges for isobaric segments or modifications with varying potential for misidentification depending on the mass spectrometer used. Additionally, the high cost and technical exper-



[labmanager.com/bp-quantumsi](http://labmanager.com/bp-quantumsi)

Explore our full Big Picture series on **proteomics!**

*Advances in Protein Identification and Sequencing* features editorial on:

- New opportunities in proteomics
- Why multiomics should include more than genomics and transcriptomics
- Strengths and limitations of current protein ID and sequencing methods



Get 3 free app notes on PTM detection, MS comparisons, and sequencing from gels!



▲ The mini-reactor wells contain a peptide tethered to its base.

tise required for analysis limits accessibility for many labs. Affinity-based assays utilize the specific interactions between a protein of interest and a complementary binding partner to capture and analyze proteins with a potentially high degree of sensitivity. However, this approach only targets expected proteins and is subject to the usual binding and detection complications, including low binding specificity, conformational changes that inhibit binding, and non-specific binding to unexpected proteins. Protein sequencing methods like Edman degradation are sensitive to amino acid resolution, but they involve lengthy processes. New and accessible technologies are needed to directly assess and identify a protein's sequence, variants, and modifications.

## NEXT-GENERATION PROTEIN SEQUENCING

The introduction of a benchtop next-generation protein sequencing (NGPS) platform enables direct sequencing of proteins from biological samples with single-molecule resolution. Though the underlying sequencing technology is entirely novel for proteins, scientists with knowledge of well-known DNA-based next-generation sequencing (NGS) platforms will find familiar principles and features. The new technology uses semiconductor chips with “mini-reactor” wells in which individual peptide molecules are immobilized, similar to ion semiconductor sequencing, pyrosequencing, and HiFi sequencing technologies for DNA. But in contrast to most DNA sequencing platforms, the NGPS technology bypasses the reverse complementarity limitation by sequencing from the N-terminal end—the opposite direction of sequencing by synthesis—trimming terminal amino acids as they are read to expose the next one for sequencing.

## HOW IT WORKS

NGPS sequences individual peptides within the integrated semiconductor chip with nanosecond precision using photosensors, optical waveguide circuitry, and reaction chambers for biomolecule immobilization. This allows for the differentiation and identification of amino acid residues, variants, and modifications based on fluorescence, lifetime, intensity, and kinetic binding pattern measurements.

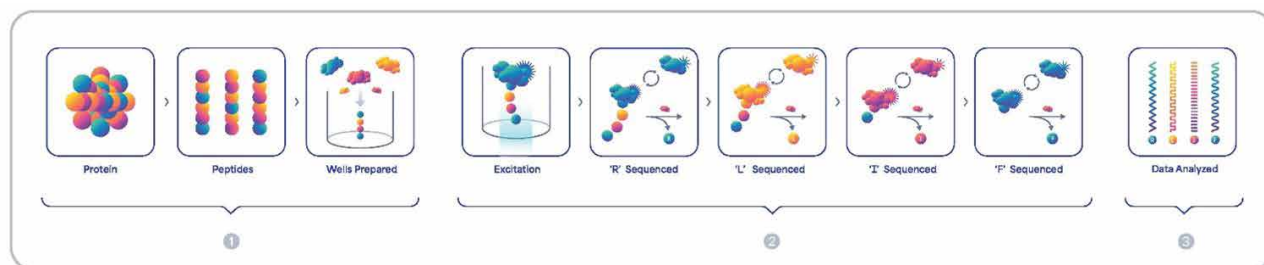
### Library preparation

Library preparation for NGPS, like in most traditional NGS, is a crucial first step and can be completed with less than three hours of hands-on time. The process begins with protein digestion into peptides. Specifically, the disulfide bonds are reduced, the cysteine residues are capped to inhibit bond reformation, and the protein is left to undergo overnight protease digestion. After that, a linker is conjugated to the peptide's side chain, creating an activated peptide-linker complex that will be immobilized on the semiconductor chip.

### Sequencing

Sequencing runs are programmed using a few simple prompts, leaving users hands-free while the instrument generates sequencing data.

During sequencing, fluorescently labeled recognizers bind the N-terminal amino acid, where they are excited by a single-wavelength pulsed laser source. They continue to bind and release from the peptide tens to hundreds of times, creating pulsing patterns characteristic of specific amino acids. These recognizers can bind more than one amino acid with differing affinities, reducing sequencing complexity by avoiding the need for specific affinity-based reagents designed against each



▲ General workflow of protein-based next-generation sequencing with Platinum® from Quantum-Si involves three phases: 1) prepare, 2) sequence, and 3) analyze.

target. Freely diffusing aminopeptidases will then cleave the N-terminal amino acid, exposing the subsequent amino acid for recognition.

Single amino acids are detected through binding kinetics, lifetime, and intensity, which are collected in real-time and securely sent to the cloud for analysis. This method is also highly sensitive to protein variants and modifications.

### Analysis

The upload of sequencing data to the cloud allows for secure access from any computer, facilitating collaboration with researchers across the globe. Data can then be analyzed using proprietary algorithms that identify trace characteristics of each binding event to determine the peptide sequence. The sequence can then be mapped to a specific protein to identify variants, mutations, or modifications. Data interpretation is automated to streamline and accelerate analytics without complex computing environments and analysis pipelines, providing multiple ways to view the data, including visualizing amino acid and peptide identification, protein mapping, confidence in calls, and other key analysis metrics.

Platinum® from Quantum-Si is currently the only NGPS platform, and it is streamlining the protein sequencing process to empower scientists with varying levels and areas of expertise. Platinum eliminates the need for highly specialized bioinformaticians churning through data, enabling any person in any lab to interpret the generated results.

### APPLICATIONS

The low cost and accessibility of this NGPS platform, combined with its ease of use, convenient benchtop size, and packaged analytics, make it an easy addition to any lab as an entry point to proteomics or for complementing other

methodologies with deeper proteomic insights. While users can identify proteins and map them to the proteome based on the predicted kinetic signatures of peptides, the most exciting aspect of this technology is its potential for discovery.

Future developments in NGPS technology have the potential to expedite disease research and diagnostics by enabling the identification of novel disease-related biomarkers and drug targets. These new proteomic insights could pave the way for improved diagnosis, monitoring, and treatment of disease and more personalized approaches to medicine. A deeper understanding of protein biology, including insights into associated proteins and their proteoforms, could enable the identification and elucidation of new mechanisms of action, leading to novel approaches for therapeutic development. Even simple analyses like identifying a protein in an SDS-PAGE band or validating the specificity of an antibody against its target protein will provide enormous value.

NGPS stands to revolutionize the fields of proteomics and multiomics as NGS did for genomics, greatly enhancing our ability to study and understand complex biological systems. Its greater accessibility due to reduced costs, small footprint, and ease of use, all within a shortened timeframe, set the stage for a similar transformation.

# QUANTUM SI™

To learn more, visit:  
[www.quantum-si.com](http://www.quantum-si.com)