QuantumSi

TECH NOTE

Protein Identification using Next-Generation Protein Sequencing of In-Gel Digested Proteins

Summary

Western blot is one of the most widely used methods for protein identification but achieving reliable and reproducible results depends on quality antibodies with validated specificity for the target protein. Furthermore, differences in amino acid sequences and isoforms cannot be easily differentiated by western blot if the size of the protein is similar. To overcome these limitations, non-antibody-based approaches are highly desired. In this technical note, we describe an in-gel protein digestion and peptide extraction procedure compatible with Quantum-Si's library preparation and protein sequencing workflows. This method can be used for the separation and enrichment of single proteins from complex mixtures using SDS-PAGE, followed by their identification on Quantum-Si's Platinum™ next-generation protein sequencing instrument.

Q-Si Technology

Quantum-Si's benchtop Platinum™ instrument enables protein sequencing from biological samples in a simple user-friendly workflow. Our technology utilizes dye-tagged N-terminal amino acid recognizers and semiconductor chip technology to detect the binding characteristics and binding order of N-terminal amino acids, resulting in unique kinetic signatures that can be used to differentiate and identify amino acid residues and PTMs. A more detailed overview of the workflow and technology can be found in our Science Paper.

Introduction

The separation and enrichment of complex samples is a key component of proteomic workflows. A relatively straightforward method for sample enrichment is gel electrophoresis. Among the different gel electrophoresis techniques, SDS-PAGE (sodium dodecyl-sulfate polyacrylamide gel electrophoresis) is widely used for protein mixtures. During this process, an electric field is applied to denatured proteins bound to the anionic detergent sodium-dodecyl sulfate (SDS), resulting in their migration through a cross-linked polyacrylamide gel matrix. This migration separates the proteins by their length and mass-to-charge ratio, resulting in individual protein bands that are resolved in the gel upon staining. While SDS-PAGE enables the separation and enrichment of complex samples, additional downstream procedures are required to identify the separated components.

Western blot is one of the most common detection methods to identify specific proteins following their separation by SDS-PAGE. In this technique, proteins are transferred from the gel to a nitrocellulose membrane, followed by blocking and then by incubation with a primary antibody specific to the protein of interest. A secondary enzyme-conjugated antibody is then used to bind to the primary antibody, resulting in a signal that can be visualized on a western blot imager for protein identification. Despite being the most frequently used antibody-based method, the availability of properly validated antibodies without off-target binding remains an issue, leading to wasted reagents, time, and money as well as erroneous data.[1] Therefore, finding alternative antibody-free methods for sensitive protein identification can enhance research productivity and reduce waste.

One promising approach for identifying pre-separated components without the use of antibodies is through next-generation protein sequencing using Quantum-Si's Platinum workflow.[2] This innovative method entails cutting proteins from the SDS-PAGE gel, digesting proteins into peptide fragments, and linking them to macromolecular linkers at the C-terminus. The peptides are then immobilized on Quantum-Si's semiconductor chip and then probed with dye-labeled N-terminal amino acid (NAA) recognizers, resulting in distinctive pulsing patterns with characteristic kinetic properties. Individual NAAs are removed sequentially by aminopeptidases in solution, uncovering subsequent amino acids for recognition. Fluorescence lifetime, intensity, and kinetic data are collected in real time and analyzed with Cloud-based software to identify the peptide sequence and corresponding protein.

To combine the benefits of separating complex mixtures by gel electrophoresis with the protein identification capabilities by Platinum, we have developed an in-gel digestion procedure that is compatible with Quantum-Si's downstream library preparation and sequencing workflows. The in-gel digestion protocol described here is similar to procedures used in other proteomic workflows.[3] To demonstrate the utility of this procedure, we prepared in-gel digested peptide libraries using the protein CDNF (Cerebral Dopamine Neurotrophic Factor) as a model protein. Sequencing of the in-gel digested peptide libraries produced single molecule peptide traces that aligned with peptides from CDNF, enabling the successful identification of the protein. These results demonstrate the compatibility of the in-gel digestion procedure with Quantum-Si's library preparation and sequencing workflow and suggest its potential in separating and enriching more complex sample mixtures to resolve differences in amino acid sequence.

Methods

We began by resuspending recombinant CDNF protein (Bio-Techne Cat # 5097-CD-050) as a 1 mg/mL solution in 125 mM HEPES, pH 8.0. Samples containing either 0.1 µg, 0.5 µg, 1 µg, 5 µg, or 10 µg of CDNF were prepared in SDS loading buffer with reductant, heated at 95°C for 5 min, and loaded onto a 4–20% Novex TGX gel (Thermo Fisher Scientific Cat # XPO4200BOX). The gel was run in Novex TGX SDS running buffer (Thermo Fisher Scientific Cat # LC2675) at 225 V for 35 min. To visualize gel bands, the staining and destaining procedure was performed with an automated eStain Coomassie staining system (Genscript Cat # L00657).

Next, we excised each band from the gel with a new razor blade and placed the gel bands on a clean surface. Each gel band was then diced into 1-mm cubes and placed into a clean 1.5-mL tube. After dicing the gel bands, we used an in-gel digestion kit (Thermo Fisher Scientific Cat # 89871) to digest the protein and extract the peptides from the gel pieces. Briefly, the gel pieces were destained, reduced, alkylated, and then washed as per the standard protocol. Next, we utilized the LysC endoprotease from the Quantum-Si Library Preparation Kit instead of trypsin in the digestion process to ensure compatibility with Quantum-Si library preparation workflow. The digestion solution was prepared by adding 1 µg of LysC endoprotease into 50 µL of the digestion buffer. We then added the full 50 µL LysC digestion solution to the dehydrated gel pieces, wrapped the tubes in parafilm, and incubated the tubes overnight (16 hrs) at 37°C with light shaking on a thermomixer. On the following day, we collected the supernatants by aspirating the solution containing the CDNF digested peptides and placed solution from each sample in a new collection tube.

To maximize recovery of the digested peptide fragments for each sample, we performed two rounds of peptide extraction on the remaining gel pieces. In the first round, 50 µL of extraction solution 1 (50% acetonitrile, 0.02% trifluoroacetic acid) was added to the tube containing the gel pieces and incubated at room temperature for 45 min. The tubes containing the gel pieces were subsequently sonicated for 5 min in a water bath, and the extraction solution was removed and placed in the tube containing the digested peptide supernatant. This process was repeated a second time but with extraction solution B (90% acetonitrile, 0.02% trifluoroacetic acid). After all gel extractions were completed, the combined peptide supernatant samples were evaporated to dryness by SpeedVac at 30°C for 60–90 minutes to remove the volatile ammonium bicarbonate from the samples and facilitate Quantum-Si's downstream library preparation workflow.

To generate the peptide library for sequencing, the digested peptides were prepared using Quantum-Si's Library Preparation Kit and Protocol with the following modification. As the peptide concentration was unknown, we resuspended the 1 µg, 5 µg, and 10 µg dried peptide samples in 50 µL sample buffer. Due to the lower input concentrations, the 0.1 µg and 0.5 µg were resuspended in 25 µL sample buffer. After preparing the digested peptide solution, we began the protocol starting at Day 2 according to the Quantum-Si Library Preparation Protocol (Document # 950-10007-00) and followed the remaining steps, except that the volume of Activators E, F, and G was reduced by a fourth or half to account for the initial 50 µL or 25 µL sample volume. To ensure an excess of peptides to Solution K during the library reaction step, all reactions contained 1 µL of Additive N and 1 µL of Solution K, however 48 µL of the derivatized peptide solution was used for the 1 µg, 5 µg, and 10 µg input samples and 23 µL of the derivatized peptide solution was used for the 0.1 µg and 0.5 µg input samples. All reactions were wrapped in parafilm and incubated overnight (16 hrs) at 37°C. On the next day, samples were removed from heat and stored on ice until sequencing on Quantum-Si's Platinum instrument using the Protein Sequencing Kit and Protocol (Document # 950-10006-00).

Results and Discussion

To explore the applicable protein input compatible with in-gel digestion in tandem with Quantum-Si's library preparation workflow, a titration series of CDNF from 0.1 µg to 10 µg was prepared. Each of the 5 samples were run side by side on a gradient gel (Figure 1), followed by excision and processing of the corresponding bands as described in the methods section (Band A). In addition to the bands at the expected CDNF molecular weight, we observed a slower migrating protein on the gel in both the 5 µg and 10 µg protein lanes (Band B). To determine the identity of this band, it was also excised from the gel of the 10 µg sample and subjected to the in-gel digestion and library preparation workflow.

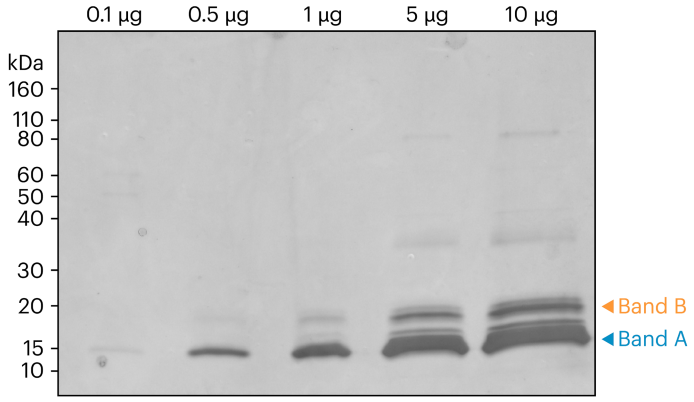

Figure 1. *SDS-PAGE Gel Image of CDNF Samples Prior to Band Excision. Orange and blue arrows indicate the approximate location of the bands removed for in-gel digestion and subsequent sequencing.*

We next sequenced the protein libraries on the Platinum instrument and generated single molecule sequencing traces, resulting in the identification of the input sample as CDNF. A total of 5 CDNF peptides were identified from the isolated 5 µg and 10 µg gel slices. Representative single molecule traces from 5 peptides are shown in Figure 2. While both the 5 µg and 10 µg libraries produced sufficient alignments to identify CDNF, the peptide libraries prepared from the 0.1 µg, 0.5 µg and 1 µg gel slices produced very few CDNF alignments (Figure 3). This outcome was expected given that the total protein starting input was outside of the recommended range of Quantum-Si's library preparation protocol, considering the potential losses from the in-gel protein digestion and peptide recovery process. This issue could potentially be mitigated in future experiments by adding a digested carrier protein to the in-gel digested peptides to ensure that the minimum protein concentration requirements are met for the library reaction.

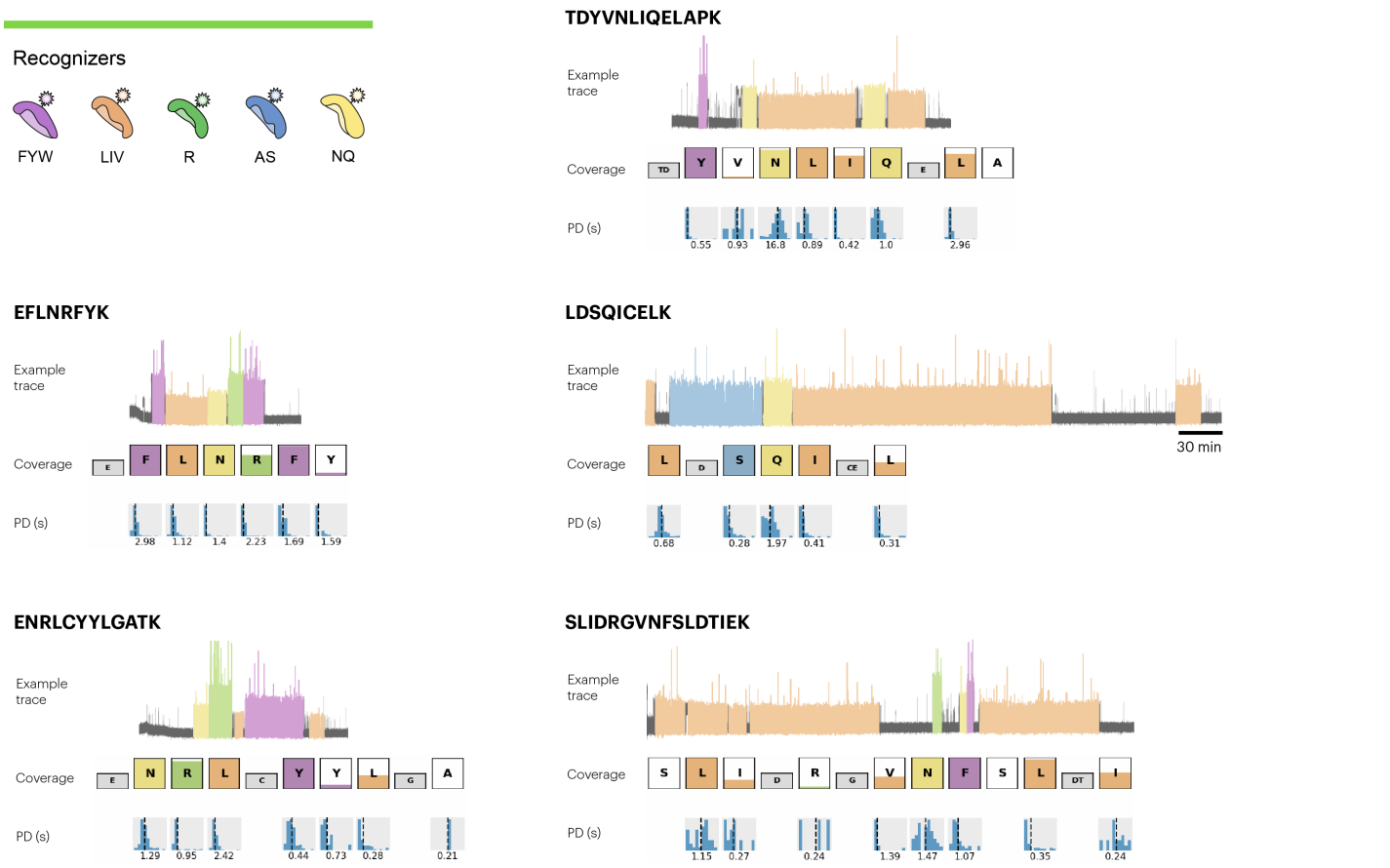

Figure 2. *Representative Traces, Coverage, and Pulse Duration Data for the Peptides Identified in the 10 µg Sample of CDNF. Five recognizers were used to identify 12 amino acids (F, Y, W, L, I, V, A, S, N, Q, R, and K), enabling identification of 5 peptides as shown in the figure.*

We also performed sequencing on the slower migrating band B in the 10 µg sample. Interestingly, this library also produced a significant number of CDNF peptide alignments, although fewer than both the 5 µg and 10 µg libraries (Figure 3). Single molecule traces representing all CDNF peptides, except ELISFCLTDK, were identified in the slower migrating 10 µg band. While the cause for the difference in the migration speed within the sample is unclear, the sequencing data clearly indicated that the slower migrating band was also CDNF.

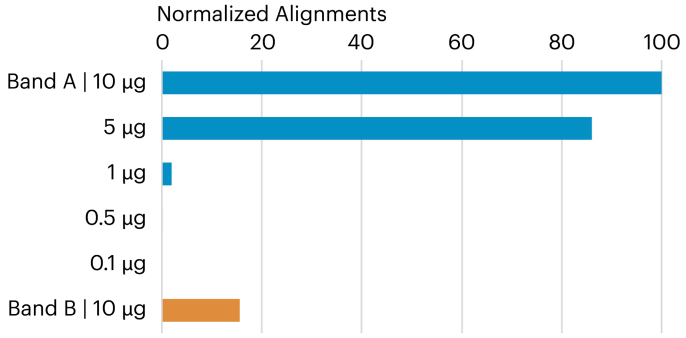

Figure 3. *Bar Graph Showing the Relative Number of Peptide Alignments for Each In-Gel Digested Sample. For Band A, both the 5 µg and 10 µg samples produced more aligned CDNF peptides compared to the 0.1 µg, 0.5 µg, and 1 µg samples. Additionally, the 10 µg Band B sample also produced a significant number of CDNF aligned peptides. All alignments normalized to the total alignments from the 10 µg library.*

Conclusion

In this technical note, we presented an in-gel digestion protocol compatible with Quantum-Si's next-generation protein sequencing workflow on Platinum. Using this procedure, we demonstrated the acceptable range of sample input for libraries prepared from in-gel digested samples that enabled the successful sequencing and identification of CDNF. This procedure has the potential to facilitate the pre-separation and identification of multiple proteins from complex mixtures, offering an alternative method to antibody-based western blot methods for protein identification from SDS-PAGE gels.

REFERENCES

1. Baker M. Reproducibility crisis: Blame it on the antibodies. *Nature*. 2015;521(7552):274-276. doi:10.1038/521274a
2. Reed BD, Meyer MJ, Abramzon V, et al. Real-time dynamic single-molecule protein sequencing on an integrated semiconductor device. *Science*. 2022;378(6616):186-192. doi:10.1126/science.abo7651
3. Shevchenko A, Tomas H, Havlis J, Olsen JV, Mann M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat Protoc*. 2006;1(6):2856-2860. doi:10.1038/nprot.2006.468