A major step forward toward high-resolution nanopore sequencing of full-length proteins

Choi, Moon Hyeok; Joo, Chirlmin

DOI
10.1016/j.molcel.2024.10.038

Publication date
2024

Document Version
Final published version

Published in
Molecular Cell

CellPress

Molecular Cell

Spotlight

A major step forward toward high-resolution nanopore sequencing of full-length proteins

Moon Hyeok Choi[1] and Chirlmin Joo[1,2,*]
[1]Department of Bionanoscience, Kavli Institute of Nanoscience, Delft University of Technology, Delft, the Netherlands
[2]Department of Physics, Ewha Womans University, Seoul, Republic of Korea
*Correspondence: c.joo@tudelft.nl
https://doi.org/10.1016/j.molcel.2024.10.038

In a recent publication in *Nature*, Motone et al.[1] report the development of a protein sequencing method using nanopores that enables the reading of long protein strands. This method allows for multi-pass re-reading and can detect single amino acid substitutions as well as post-translational modifications (PTMs).

Developing single-molecule protein sequencing technologies is crucial to fully understanding the complexity of the proteome, as proteins, the workhorses of cellular processes, exist in many different forms called proteoforms.[2] The proteome is dynamic and highly diverse, making it essential to decode proteins at the single-molecule level to identify rare proteoforms, especially those involved in disease. Current techniques often miss this heterogeneity, leaving a large portion of the proteome uncharacterized. Emerging single-molecule methods such as nanopore-based approaches[3] and fluorescence-based techniques[4] are promising solutions for protein sequencing. In particular, full-length protein sequencing, which enables the identification of both amino acid sequences and post-translational modifications (PTMs), could revolutionize our understanding of biology and open new frontiers in personalized medicine, diagnostics, and therapeutic development.[5]

In the study by Motone et al.,[1] the authors achieved a major breakthrough by demonstrating the ability to sequence, using nanopore technology, significantly longer protein strands than was previously possible. By combining the CsgG nanopore with the ClpX unfoldase, they unfolded and translocated long protein sequences through the nanopore, overcoming a key limitation of earlier techniques, which faced challenges with long peptides (Figure 1). In earlier studies, the basic principle of protein translocation through nanopores was introduced using α-hemolysin nanopores, and the major hurdles involved the irregular motion of proteins through the pore, which lowered resolution and made reliable sequencing difficult (Figure 1).[6] Although ClpX unfoldase was already employed in those studies to unfold and translocate proteins, the ability to detect and identify single amino acids was limited by the lack of precise control over the translocation speed and the difficulty in maintaining the protein's unfolded state throughout the process.[7] In contrast, this study introduces a refined use of CsgG nanopores and a better-controlled application of the ClpX unfoldase, allowing for the precise stepwise movement of proteins through the nanopore while maintaining high sensitivity to amino acid variations and PTMs.

What sets this method apart from the previous ClpX-based strategies is the motor-driven, reverse-direction pull using ClpX, which provides much greater control compared to the voltage-driven translocation in earlier approaches. While ClpX pulls proteins in two amino acid steps regardless of direction, the reverse-direction pull ensures smoother, stepwise movement through the pore, producing cleaner, higher-resolution signals. As a result, this system offers more precise detection of amino acid sequences and PTMs compared to earlier methods.[9] Another standout feature of this study is the multi-pass reading capability, which allowed the same protein strand to be read multiple times as it moved in and out of the pore, further enhancing the accuracy as any potential errors from a single pass can be corrected through subsequent reads (Figure 1).[8]

It should be noted that recent nanopore protein sequencing methods relied on conjugating oligonucleotides to peptides to control translocation and read the peptide sequences.[8–10] Enzymes that recognize these oligonucleotides were critical in the movement of peptides in a controlled manner. However, this method had limitations in terms of the length of peptides that could be measured, making it ineffective for reading long protein strands. In this study,[1] the authors demonstrated that the ClpX-mediated translocation method solves the issue. Unlike oligonucleotide-based methods, ClpX translocates proteins, ensuring consistent, long movement through the nanopore, which therefore allows for the sequencing of full-length molecules that were previously unattainable using nanopores.

Despite the significant advancements achieved in this study, there remain a few important areas where further development could be beneficial. One of the main challenges is that the technology has primarily shown success with designed protein sequences, which are engineered for easier translocation and unfolding. This does not fully capture the natural complexity and diversity of proteins found in biological systems. As a result, while the sequencing of designed proteins holds great promise, applying this technology to the *de novo* sequencing of naturally occurring, full-length proteins remains an ongoing challenge.

Additionally, the reliance on ClpX unfoldase introduces some limitations regarding the types of proteins that can be processed. Highly stable proteins with complex tertiary and quaternary structures may resist unfolding, making them more difficult to sequence using this

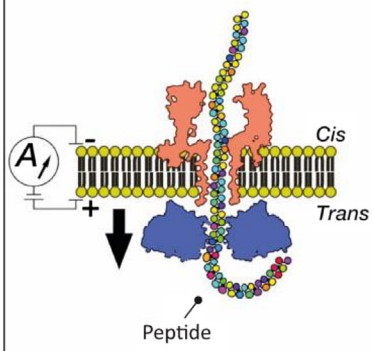
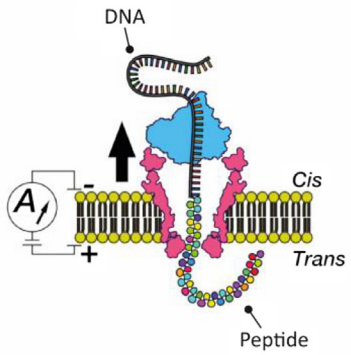
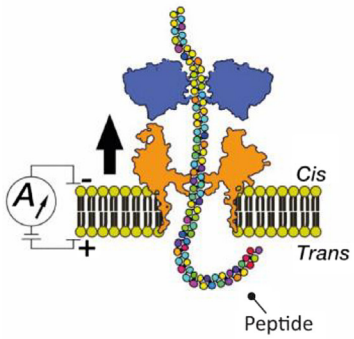
	Nivala, et al. Nat. Biotechnol. (2013)	Brinkerhoff, et al. Science (2021)	Motione, et al. Nature (2024)
Schematic Diagram			
Nanopore	**α-hemolysin**	**MspA**	**CsgG**
Translocase	**ClpX**	**Hel308**	**ClpX**
Reading Direction	***Cis* to *Trans***	***Trans* to *Cis***	***Trans* to *Cis***
Multiple Re-reading	**Not Available**	**Available**	**Available**

Figure 1. Comparison of nanopore protein sequencing technologies across different studies
The figure highlights key differences in schematic diagram, nanopore types, translocases used, reading direction, and the capability for multiple re-readings between the methods developed by Nivala et al.,[6] Brinkerhoff et al.,[8] and Motone et al.[1]

method. Furthermore, although this study successfully detected phosphorylation PTMs, it is not yet capable of identifying all types of PTMs, especially those that are not highly charged or not bulky. As such, the current nanopore sequencing approach has not fully addressed PTMs, which are crucial to understanding protein function.

Moving forward, protein sequencing technology will likely focus on enhancing its ability to process native proteins without requiring extensive modifications. Efforts will also be directed toward discovering and engineering novel unfoldases or nanopores. Along with reducing the step size to one (or less than one) amino acid, this will significantly enhance sequencing resolution. Furthermore, gaining a deeper understanding of how PTMs affect ionic currents will enable more comprehensive and accurate protein analysis. These advancements will expand the applications of nanopore technology, bringing us closer to routine single-molecule protein sequencing.

ACKNOWLEDGMENTS

C.J. acknowledges funding from NWO (Vici, VI.C.202.015), the Basic Science Research Program (NRF-2023R1A2C2004745), and Frontier 10-10 (Ewha Womans University). We thank Tor Eriksson for his critical reading.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. Motone, K., Kontogiorgos-Heintz, D., Wee, J., Kurihara, K., Yang, S., Roote, G., Fox, O.E., Fang, Y., Queen, M., Tolhurst, M., et al. (2024). Multi-pass, single-molecule nanopore reading of long protein strands. Nature *633*, 662–669. https://doi.org/10.1038/s41586-024-07935-7.

2. Alfaro, J.A., Bohländer, P., Dai, M., Filius, M., Howard, C.J., Van Kooten, X.F., Ohayon, S., Pomorski, A., Schmid, S., Aksimentiev, A., et al. (2021). The emerging landscape of single-molecule protein sequencing technologies. Nat. Methods *18*, 604–617. https://doi.org/10.1038/s41592-021-01143-1.

3. Hu, Z.L., Huo, M.Z., Ying, Y.L., and Long, Y.T. (2021). Biological nanopore approach for single-molecule protein sequencing. A. Chem. Int., ed. *60*, 14738–14749. https://doi.org/10.1002/anie.202013462.

4. Floyd, B.M., and Marcotte, E.M. (2022). Protein sequencing, one molecule at a time. Annu. Rev. Biophys. *51*, 181–200. https://doi.org/10.1146/annurev-biophys-102121-103615.

5. Filius, M., van Wee, R., de Lannoy, C., Westerlaken, I., Li, Z., Kim, S.H., de Agrela Pinto, C., Wu, Y., Boons, G.J., Pabst, M., et al. (2024). Full-length single-molecule protein fingerprinting. Nat. Nanotechnol. *19*, 652–659. https://doi.org/10.1038/s41565-023-01598-7.

6. Nivala, J., Marks, D.B., and Akeson, M. (2013). Unfoldase-mediated protein translocation through an α-hemolysin nanopore. Nat. Biotechnol. *31*, 247–250. https://doi.org/10.1038/nbt.2503.

7. Sen, M., Maillard, R.A., Nyquist, K., Rodriguez-Aliaga, P., Pressé, S., Martin, A., and Bustamante, C. (2013). The ClpXP protease unfolds substrates using a constant rate of pulling but different gears. Cell *155*, 636–646. https://doi.org/10.1016/j.cell.2013.09.022.

8. Brinkerhoff, H., Kang, A.S.W., Liu, J., Aksimentiev, A., and Dekker, C. (2021). Multiple rereads of single proteins at single–amino acid resolution using nanopores. Science *374*, 1509–1513. https://doi.org/10.1126/science.abl4381.

9. Nova, I.C., Ritmejeris, J., Brinkerhoff, H., Koenig, T.J.R., Gundlach, J.H., and Dekker, C. (2024). Detection of phosphorylation post-translational modifications along single peptides with nanopores. Nat. Biotechnol. *42*, 710–714. https://doi.org/10.1038/s41587-023-01839-z.

10. Yan, S., Zhang, J., Wang, Y., Guo, W., Zhang, S., Liu, Y., Cao, J., Wang, Y., Wang, L., Ma, F., et al. (2021). Single Molecule Ratcheting Motion of Peptides in a Mycobacterium smegmatis Porin A (MspA) Nanopore. Nano Lett. *21*, 6703–6710. https://doi.org/10.1021/acs.nanolett.1c02371.