Nanopore sequencing

https://doi.org/10.1038/s41565-023-01480-6

Unfolding the path to nanopore protein sequencing

Adam Dorey & Stefan Howorka

Check for updates

A modified nanopore enables enzyme-free threading of single polypeptides to detect post-translational modifications

DNA sequencing with nanopores has had a considerable impact during the COVID-19 pandemic by identifying SARS-CoV-2 variants fast. More generally, nanopores provide step-changing portable and long-read sequencing[1,2], simply by threading DNA strands through ring-like nanopores. Can nanopores also be used for protein sequencing? This is challenging, as the compactly folded proteins thread less easily through a nanopore than elongated DNA. Now writing in *Nature Nanotechnology*, Martin-Baniandres et al.[3] describe how protein unravelling and threading helps determine the proteins' biologically relevant post-translational modifications (PTMs) at single-molecule resolution.

In classical nanopore sequencing, individual single-stranded DNA molecules of negative charge are electrophoretically transported through a 1 nm-wide lumen of a protein pore[2,4,5]. Transport is additionally controlled with a molecular motor enzyme to feed the DNA strand base-by-base into the lumen[4]. The sequence is read as each base causes a distinct electrical blockade in the ionic current flowing across the pore. Replicating electrical sequencing for proteins would be a critical scientific and technological advance. But protein sequencing is no easy feat. Proteins are not only more compactly folded than DNA; deciphering 20 amino acids is also more challenging than reading four nucleobases. Furthermore, mature proteins contain chemically complex PTMs.

To turn protein sequencing with nanopores into reality, researchers must achieve two aims: resolving the 20 amino acid residues and their PTMs, and controlling polypeptide unfolding and threading through the pore. Regarding the first aim, 20 amino acids could be discriminated when single residues were attached to a carrier peptide[6]. Clearly, nanopores are sensitive enough to distinguish the often minuscule differences in residue size and structure.

Relating to the second aim, Martin-Baniandres et al.[3] advance the field by threading proteins across a nanopore and detecting PTMs within the unfolded polypeptide chain (Fig. 1). Crucially, the authors use electro-osmosis (EO) to thread proteins. The traditional electrophoretic pulling as used for negatively charged DNA does not work for polypeptides due to their charge-neutral backbone and the heterogeneously charged residues. In alternative EO, an electric field drives the flux of buffer solution, which co-transports the protein chain through the nanopore.

Martin-Baniandres et al.[3] achieve EO with an engineered protein nanopore carrying a ring of positively charged amino acids at its narrowest part. This electrostatic filter causes strong flux of electrolyte anions and the surrounding water, which co-transports the polypeptide chain. Protein unfolding and threading is enhanced with protein-denaturing agent guanidium hydrochloride (GdHCl), which is added to the buffer at a concentration of 0.75 M. The authors

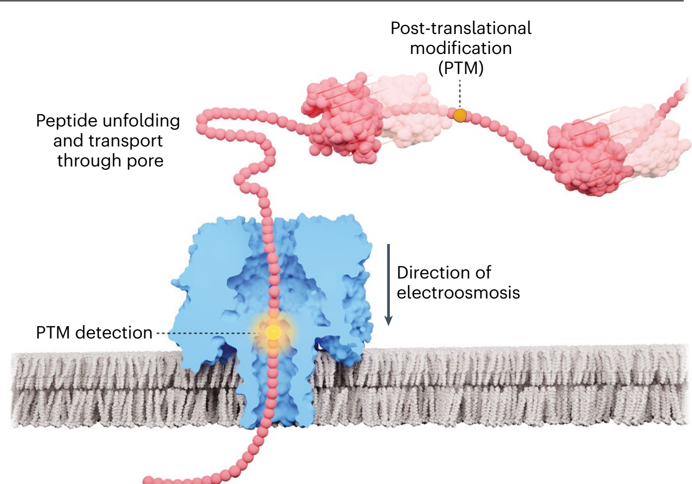

Fig. 1 | Protein threading through a nanopore detects post-translational modifications. A polypeptide (pink) with globular repeats is driven through a protein nanopore (blue) embedded in a bilayer membrane (grey). Electro-osmosis drives the threading and helps detect PTMs (orange, yellow when in the nanopore) in the polypeptide.

demonstrate EO-mediated threading with analyte proteins engineered to carry multiple interlinked globular repeats (Fig. 1). As threading of a single repeat causes a characteristic unfolding signature, multiple corresponding signatures can be expected for proteins with 2, 4, 6 and 8 repeats. Indeed, the read-outs largely show the pattern, even though a few repeats are missed.

To determine whether PTMs can also be detected, the authors use a protein with a single modification in between two repeats (Fig. 1). Three modification types are tested: phosphorylation, glycosylation, and glutathionylation, which are among the most common PTMs. Upon pore threading, each modification causes a reduction in electrical signal proportional to the mass of the PTM. Significantly, the resolution is high enough to discriminate between a mixture of proteins with different PTMs.

How practical is this approach, which is shown for one engineered protein, to proteomic analysis? As native proteins vary significantly in their resistance to unfolding, universal conditions do not exist to capture and thread all proteins within a cell. A potential solution is to utilise voltage sweeps in combination with denaturants to cover most proteins. Furthermore, insufficient signatures of very small PTMs may be amplified with cognate antibodies or chemical binders.

How does the approach relate to other techniques? Electro-osmosis has been used previously for polypeptide transport through a nanopore lacking the electrostatic filter. In contrast, EO was induced with a higher GdHCl concentration of up to 2 M as the

charged agent can adsorb to the lumen's surface and thereby induce the EO-characteristic flow of counterions and water (ref. 7). Using this system, longer proteins threaded at 10 µs per residue which is slower than in electrophoresis and may help improve resolution. However, the sensing of PTMs was not tested.

Alternatively, protein threading can be driven by enzymes. For example, a nanomachine was engineered by fusing a ring-like unfoldase molecular motor or a protease onto a nanopore[8]. The enzymatic activities allowed for either threading long polypeptides or shorter peptides from protease-fragmentation, respectively. The strands were detected, yet individual amino acids were not resolved.

Another enzyme-driven method did, however, resolve single amino acids in a short peptide by cleverly harnessing classical nanopore DNA sequencing with a molecular motor[9]. As the peptide was fused to a carrier oligonucleotide, the molecular motor not only threaded the DNA but also the linked peptide strand to allow reading of sequential residues. Re-threading the peptide strand increased sequencing accuracies to over 99.99%. This is a significant leap forward but currently limited to peptides <25 residues in length and peptides of predominantly negative charge.

So far, the market for protein sequencing is dominated by Edman degradation and mass spectrometry. With these techniques, peptides with 30–40 residues are routinely sequenced to correctly predict the proteins' identity from a database and to detect the position and nature of PTMs. While proteins are detected at picomole to femtomole levels, the sensitivity is insufficient for analysing the proteome of single cells, which is common for genomic DNA and RNA sequencing. Single-cell proteome analysis is an area where single-molecule methods can contribute[10].

It is currently difficult to predict which nanopore technology, whether enzyme-driven, enzyme-free, or combinations thereof, will come to market for protein sequencing. Techniques will probably be tailored for applications where analysis parameters such as ease of preparation, accuracy, cost-effectiveness, read length, portability, and scalability are matched to the specific demand. These developments will also take advantage of existing DNA sequencing platforms and new commercial entrants.

Once, nanopore DNA sequencing was regarded as an outlandish idea and deemed impossible by some leading scientists. Yet, interdisciplinary research and iterative improvements — underpinned by a clear vision and dogged determination — have yielded this powerful, commercially available, and widely used sequencing platform. We are now in a similar phase of exciting technological developments for nanopore protein sequencing and other single-molecule approaches[11].

Adam Dorey & Stefan Howorka
Department of Chemistry, Institute of Structural and Molecular Biology, University College London, London, UK.
e-mail: s.howorka@ucl.ac.uk

Published online: 25 August 2023

References

1. Marx, V. *Nat. Methods* **20**, 6–11 (2023).
2. Quick, J. et al. *Nature* **530**, 228–232 (2016).
3. Martin-Baniandres, P. et al. *Nat. Nanotechnol.* https://doi.org/10.1038/s41565-023-01462-8 (2023).
4. Wang, Y. Y., Zhao, Y., Bollas, A., Wang, Y. Y. & Au, K. F. *Nat. Biotechnol.* **39**, 1348–1365 (2021).
5. Goyal, P. et al. *Nature* **516**, 250–253 (2014).
6. Ouldali, H. et al. *Nat. Biotechnol.* **38**, 176–181 (2020).
7. Yu, L. et al. *Nat. Biotechnol.* **41**, 1130–1139 (2023).
8. Zhang, S. et al. *Nat. Chem.* **13**, 1192–1199 (2021).
9. Brinkerhoff, H., Kang, A. S. W., Liu, J., Aksimentiev, A. & Dekker, C. *Science* **374**, 1509–1513 (2021).
10. Reed, B. D. et al. *Science* **378**, 186–192 (2022).
11. Alfaro, J. A. et al. *Nat. Methods* **18**, 604–617 (2021).

Competing interests

Stefan Howorka is inventor on a patent on a protein pore used by Oxford Nanopore Technologies plc for DNA sequencing.