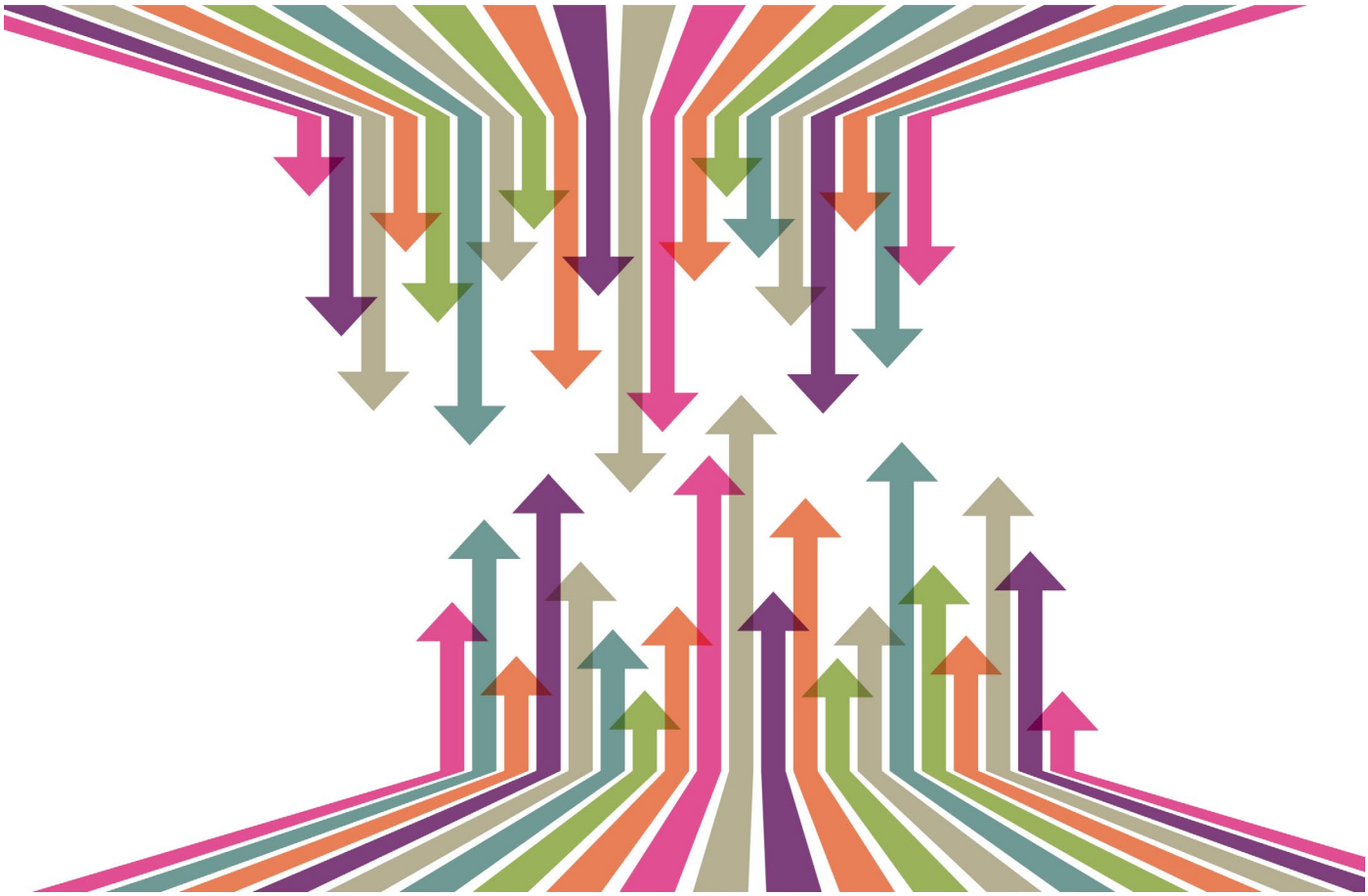
https://doi.org/10.1038/s41592-024-02170-4



Inside the chase after those elusive proteoforms

Human cells contain crowds of protein variants, but, especially in a time of funding challenges, chasing these proteoforms takes dogged persistence.

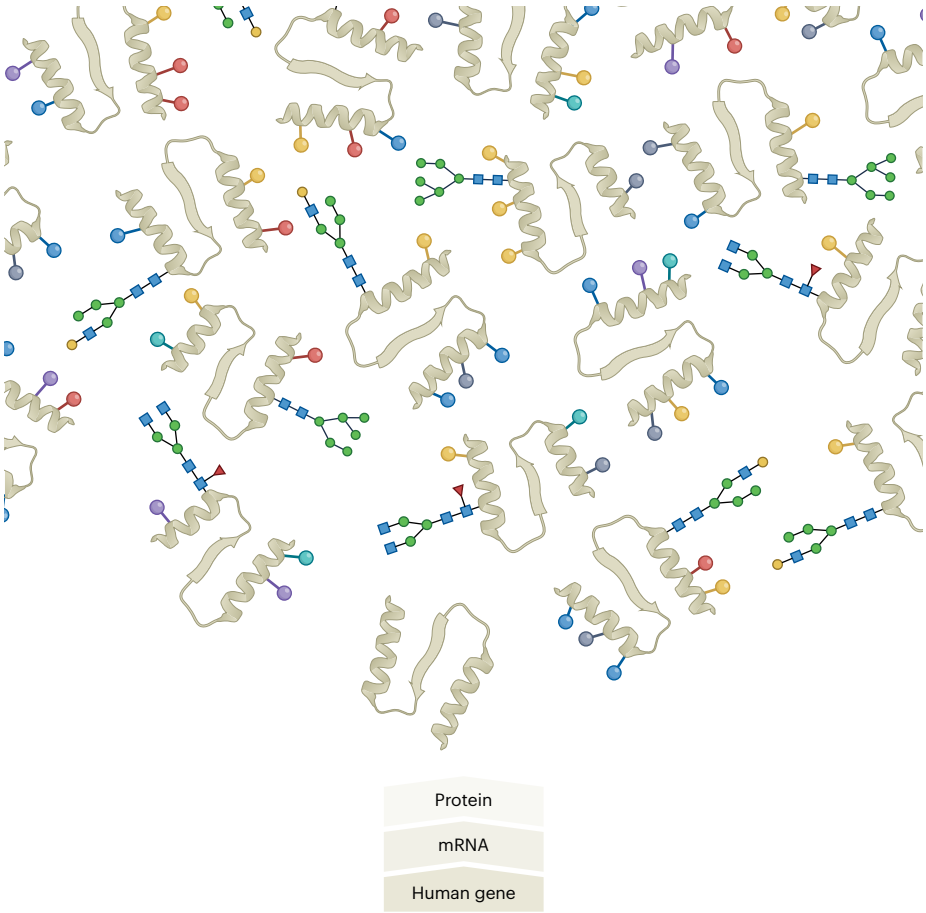
By Vivien Marx

Biology delivers a massive number of puzzles to proteoform hunters[1,2]. Proteoforms are the droves of protein variants that one and the same gene can give rise to. This variant explosion takes place after DNA is transcribed to mRNA and a protein is synthesized. No single method currently lets researchers hunt these proteoforms reliably and at scale. Sparks fly in discussions about which of the existing approaches — 'bottom-up' or 'top-down' proteomics, shades thereof, or other, newer technologies — is most promising. Some rifts seem to be closing as scientists keep up their proteoform hunt in a tense funding environment.

Parsing the multitude of protein variants, says University of Liverpool researcher Claire Eyers, would open up a wealth of poorly understood biology. 'Proteoform' describes the actual functional protein in a cell, "rather than what you think may be expressed based on the gene," says Eyers. Kiel University researcher Andreas Tholey would relish it if proteomics became proteoform-centric. The classic one-gene, one-protein, one-function view yields too many misinterpretations about the proteome.

Lloyd Smith of the University of Wisconsin, Neil Kelleher from Northwestern University and others in the Consortium for Top-Down Proteomics introduced[3] the term 'proteoform' in 2013 as one that encompasses "all of the

CREDIT: F. COELHO / GETTY IMAGES



Once DNA is transcribed to mRNA and a protein is synthesized, variants can multiply explosively as a result of the many different types of post-translational changes to a protein.

different molecular forms in which the protein product of a single gene can be found," which include changes due to genetic variation, splicing and post-translational modifications.

University of Washington researcher Michael MacCoss says he and Kelleher have occasionally "bumped heads" about proteoforms and methods to assess them. "I'm a bottom-up person," says MacCoss. They agree that proteoforms matter, as do strategies about how best to find and characterize them. "To me, proteoforms tell the whole story of a protein," says Benjamin Garcia from Washington University School of Medicine in St. Louis, who is the president-elect of the US Human Proteome Organization.

What's it called?

The word proteoform has gained acceptance in the proteomics community and "would be a good addition to the biochemistry and biology dictionary," says Salvatore Sechi, who is the program director for proteomics and systems biology at the US National Institutes of Health (NIH) National Institute of Diabetes and Digestive and Kidney Diseases. John Yates from Scripps Research Institute sees increased acceptance, too, but, he says, "I'm not sure how far out of proteomics it has permeated." Within the lab, University of Oxford researcher Carol Robinson and her group talk about proteoforms, but they do not use the word more widely without explaining it, as in her view it is not yet fully accepted. Eyers says ever more are familiar with the term, but "I don't use it without explaining what it means." MacCoss likes the word 'proteoform' a lot, he says. Proteoforms are hard to analyze, which is part of why the proteomics field is setting them aside, he says. But it matters to measure and characterize them given their diverse functions.

A proteoform might be a protein from which a few amino acids have been cleaved, a change that alters the protein's functions in the immune system, says Kelleher. Approximately 1,000 variants exist of human hemoglobin, says Sechi. This protein in red blood cells has a proteoform called hemoglobin A1c. When, after synthesis, hemoglobin is exposed to glucose in the blood, it can become glycated hemoglobin A1c. People with prediabetes or diabetes have more glycated hemoglobin, which is why this proteoform is used as a biomarker.

Many biomarkers are proteoforms, says MacCoss. The peptide amyloid-β (Aβ) has many proteoforms. These peptides are post-translationally processed from the same transmembrane protein. Aβ40 and Aβ42 have been identified post-mortem in the brains of people who had Alzheimer's disease. Another protein associated with Alzheimer's is tau, of which there are also many variants, which can differ by phosphorylation sites. It is not entirely clear what these proteoforms do.

Some proteoforms can be artifacts that arise from the way a sample is processed, says Eyers. So one must separate those from ones that promise biologically interesting post-translational modifications (PTMs). PTMs include disulfide bonds, which stabilize protein structure, and then there are the many kinds of covalent chemical modifications such as phosphorylation or methylation. "There's a lot, and we tend to focus on a handful of those," Eyers says. The classic number of human protein PTMs, says Kelleher, hovers around 250, "and that number has gone up." When considering the biologically relevant PTMs, says Eyers, the figure she uses is around 300.

When considering the biologically relevant PTMs, says Claire Eyers, the figure she uses is around 300.

Proteins can be decorated with different combinations of PTMs. Some estimate a protein with 50 PTMs to have as many as 2^{50} proteoforms, but that's "absurdly large," says Kelleher, and an incorrect "world view." To a certain extent, the number is conjecture, says Smith who was interviewed jointly with Kelleher.

Historically, says MacCoss, researchers have assessed and counted objects of interest once they had the tools to do so. In proteomics, mass spectrometry is the tool for discerning molecules according to their mass-to-charge ratio, but it is far from straightforward. Acetylated and trimethylated protein fragments, for instance, have similar mass-to-charge ratios. The increased accuracy of mass measurement

CREDIT: T. PHILLIPS, SPRINGER NATURE. SOURCES: C. EYERS, UNIV. LIVERPOOL; B. GARCIA, WASHINGTON UNIV.; N. KELLEHER, NORTHWESTERN UNIV.; LLOYD SMITH, UNIV. WISCONSIN

Claire Eyers (back row, second from right) at the University of Liverpool and her team use bottom-up and top-down proteomics as well as native mass spec. In her view, parsing the multitude of protein variants would open up a wealth of poorly understood biology.

CREDIT: UNIV. LIVERPOOL

has made it possible to tell them apart. Further advances add ever more resolution to mass spec data. With proteoforms, he says, "at some point in time, we have to figure out a way to measure these things," reliably, robustly and precisely.

Bottom-up, top-down

In bottom-up proteomics, says Eyers, researchers chop up proteins and analyze peptides. In top- down proteomics, whole, intact proteins are analyzed. Kelleher, working with Fred McLafferty and others, has advanced top-down proteomics, which involves intact proteins[4]. In native mass spec, the idea, says Eyers, is to additionally analyze conformation dynamics of intact proteins, and some groups analyze protein complexes this way. She and her team use bottom-up and top-down proteomics, as well as native mass spec. "We actually do those three to be able to understand the effect of protein modifications on primary through to tertiary structure," says Eyers. Changes can affect how a protein binds to a small-molecule drug.

She and her team look at the roles protein kinase A proteoforms might play in Cushing's syndrome. Protein kinase A has 45 theoretical phosphorylation sites at serine and threonine residues. With bottom-up proteomics, 12 phosphorylation events can be found, and top-down proteomics software can find more. The challenge, she says, rests with understanding the complex PTM combinations.

For proteins larger than around 30 kilodaltons, says Robinson, "it is more common to chop them up into smaller, more manageable pieces for bottom-up analyses than it is to do top-down." Taking the bottom-up route with peptides, though, makes it hard, for example, to relate different PTMs, such as palmitoylation status and phospho-forms, to one another.

The Yates lab has advanced bottom-up proteomics[5], and his lab still primarily uses bottom-up proteomics for tasks such as characterizing HIV glycoprotein P120/140 to determine sites and types of glycosylation. "For this problem I probably wouldn't use top-down, as glycan heterogeneity makes top-down too hard," he says. "We are slowly moving into top-down for certain problems," he says, such as for single-cell analysis. In his view, "top-down methods will be used more generally as instruments and software become more capable." What has advanced top-down proteomics and native mass spec, in his view, are the expanded capabilities of the Orbitrap mass spectrometer.

In his view, says John Yates, "top-down methods will be used more generally as instruments and software become more capable."

As an example of how PTM patterns dictate function, says Yates, the top-down community uses modifications of the tail of the histone protein. "Histones may be an extreme case of this," he says. True, he says, top-down and native mass spec have the advantage of offering a view of the whole unit, "but basically top-down still needs to convincingly demonstrate they can discover new biology with their methods." It seems to him that top-down is "at the same stage protein mass spectrometry was 35 years ago." Back then, bottom-up was used mostly to confirm hypotheses, but "bottom-up has been more on the hypothesis-generating side of biology for the last 20 years."

Bottom-up proteomics, however, cannot readily scan all regions of a protein, which hinders the hunt for proteoforms. One can enhance coverage, says Yates. To find glycosylation sites on the SARS-CoV-2 spike protein, he and his team developed a digestion approach with Proteinase K and tweaked buffers, which helped them find many more glycans than with classic bottom-up analysis.

The level of fragmentation in top-down tandem mass spec results in low sequence coverage, says Yates, and "and one can argue that without fragment ions bracketing a modification, its site localization is ambiguous." Work ongoing in several labs is improving ion activation, which increases sequence coverage and can generalize the tandem mass spec process. With larger proteins, tandem mass spec on intact proteins is less efficient. "What distinguishes bottom-up from top-down is that peptide mass spectrometry is far more mature, robust and reproducible," he says. "But top-down advances."

It's not enough to look at a piece of a protein using bottom-up proteomics, says Garcia. "It's similar to reading a book but only reading every fifth word," he says. One can sort out meaning, but it's not like reading the entire sentence.

Kelleher sees the proteomics field divided into bottom-up or top-down, and beyond that there is non-mass spec-based analysis, which includes assays from commercial vendors such as Olink and SomaLogic. "The bottom-up people have the megaphone when it comes to the mass spectrometry," he says.

The top-down proteomics community evolves much as the bottom-up community did, says Tholey. After working out how to measure proteins and proteoforms, among next steps in top-down proteomics is developing methods to sort through false positives and quality-control the data. He and his team may have found a new proteoform, says Hartmut Schlüter from University Medical Center Hamburg-Eppendorf, who was interviewed jointly with Tholey. The result still needs to be validated and functionally assessed.

Unlike in work with genes, no method exists with which to amplify proteins. One validation approach is to find an enzyme that installs and

uninstalls the PTM of interest. "If we know that there is a very strong effect on the activity of a protein, then this can also help to identify the meaning of a post-translational modification," says Schlüter. Careful functional studies matter. Aβ undergoes many changes, and some proteoforms can indicate disease while others do not, says Tholey.

Tholey draws parallels between proteomics and astronomy. The Hubble Space Telescope helped science to move beyond the limitations of Earth-based telescopes, which led to "huge new developments" in astronomy. Progress in proteomics has taken place, and "we are now at the Hubble stage," he says. "We have to convince people that it is important to get the James Webb Telescope also for proteomics," referring to the new infrared space observatory.

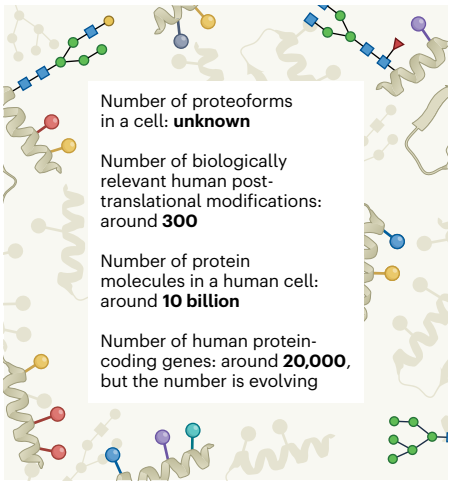
"Top-down tools are hard to use," says Lloyd Smith. "So we're working on improving that."

At the moment, says Eyers, "we're at the stage where people tend to either be in the bottom-up camp or the top-down camp." Her lab and the Yates lab at the Scripps Research Institute "kind of straddle both," she says. "I don't think that we are approaching the stage where we can do top-down proteomics properly without it being informed by bottom-up proteomics information." One can find out which protein a peptide belongs to, she says, but one must constrain the search space to defining PTM patterns and combinations. MacCoss hopes scientists will start including proteoform numbers in their measurements instead of the common practice of summarizing a single measurement per protein. One protein can exist in multiple proteoforms[6].

Computed proteoforms

As he plans his career in proteomics, Allen Po, a PhD student in the Eyers lab, sees that findings from bottom-up studies play into top-down approaches. "The top-down field is quite new still," says Po, and top-down software can make data analysis challenging. At conferences, when he mentions he studies intact proteins, people tell him "it's the future," he says. "Whether they mean it or not, I don't know."

Around one-third of the Smith lab does informatics. "Top-down tools are hard to use," says Smith. "So we're working on improving that." Analysis challenges are typical in all areas with complex biology, complex chemistry and complex instrumentation coupled with massive data.



More and more people are using top-down as the methods and computation become easier and robust, says Garcia. "It's a process, so the top-down community is still working to make the approach more accessible, but it's not quite there yet in my opinion."

Eyers and her team work on various informatics projects, such as search tools for identifying proteoforms. Taking a software tool's output at face value can lead to misinterpretations and missed data nuances. A biologist exploring which proteoforms matter in a system or a pathway needs scalable methods, she says. "I think it's not there yet." Once proteoforms have been identified, one needs to learn what each one does. "There's multiple technical challenges, which I think is prohibiting a cultural shift," she says.

An increasing number of researchers work on proteoform challenges and use different methods, says Eyers. "There's problems at every single step," she says, be that sample prep or fragmentation inside the mass spec or data analysis. Around 30% of proteins, such as transmembrane proteins, are entirely insoluble in the typical buffers used for electrospray, says MacCoss. Electrospray is how a sample gets into the mass spec instrument.

"The question I ask myself not infrequently," says Eyers, "is whether or not mass spectrometry is going to be the way that we are doing proteoform characterization in 10 years." She is working with Hagan Bayley and others at University of Oxford to develop a nanopore-based approach to distinguish proteoforms on the basis of measurements made as a protein moves through a nanopore. Such approaches are gaining traction for proteoform analysis and other questions, too, she says. One big challenge is going to be throughput and scale.

MacCoss wonders about nanopore-based approaches and scale. It's tough to apply a technology that works well for RNA sequencing to proteomics. A human cell has around 300,000 mRNA molecules and around 10 billion protein molecules, which can be modified in myriad ways. "There's just too many molecules," he says. Liquid chromatography combined with mass spec can measure billions of molecules per hour and sequencing methods take days to measure billions of reads, he says.

The new non-mass spec approaches, says Garcia, are pretty amazing. These sequencing approaches are not yet ready to replace mass spec, but "they will make a great impact in the future," he says. "When? I am not sure." Mass spec is still so flexible and versatile, he says, that many applications will be difficult to replace fully by other approaches.

Tholey also sees mass spec as the major tool for the next little while, but one must not neglect emerging technologies such as nanopore sequencing and others. Eventually, when scientists want to combine data, it may turn out to be challenging to align mass spec output and newer approaches. One method might be more suited for quality control of findings from another.

"New approaches are on the horizon. I think they will be transformative," says Carol Robinson.

Membrane proteins are difficult to analyze in that they are hydrophobic, insoluble and clump together in standard mass spec buffers, says Robinson. Her lab, in collaboration with colleagues at other institutions, studies these proteins using native mass spec. They engineer instruments and tweak buffers to keep the protein and its partners intact during electrospray ionization[7].

The divide between bottom-up and top-down proteomics exists, says Robinson. But, in her observation, "yes, definitely I think the divide is closing," and methods development for proteoform analysis stands to benefit. Top-down has, in her view, been seriously limited by software that must cope with the complexity of the datasets and the technical limitations of available instrumentation, which is something that can be pushed beyond current limits, she says. "New approaches are on the horizon. I think they will be transformative."

CREDIT: T. PHILLIPS, SPRINGER NATURE. SOURCES: C. EYERS, UNIV. LIVERPOOL; B. GARCIA, WASHINGTON UNIV.; M. MACCOSS, UNIV. WASHINGTON; H. SCHLÜTER, UNIV. MED. CTR HAMBURG-EPPENDORF; L. SMITH, UNIV. WISCONSIN; J. YATES, SCRIPPS RESEARCH INST.

Top-down of native complexes is, she says, "the ultimate dream" and would in her view involve ways to directly capture lipids, PTMs and interactions all in the "cell soup." In one set of experiments, she and her team teased out molecular details of a signaling cascade at G protein-coupled receptors[8].

Ultimately, she says, the goal is to unpack the intricate processes in our cells. "PTM status — and by extension proteoform identity — is dynamic and context-dependent," she says. Understanding the temporal changes that occur and the order of change during signaling, for example, is key to understanding the molecular underpinnings of health and disease. Says Robinson, "these dynamic cell-based changes are the major driver of a lot of new biological discovery."

Protein sequencing and proteoforms

Each type of protein exists in a wide range of highly dynamic structures that are governed by splice variants and post-translational modifications, says Stanford University researcher Parag Mallick. Proteoforms may differ between tissue types or disease. Much points to molecular heterogeneity of proteoforms being hugely important, but to date it's been hard to measure. "Whenever there is a measurement gap, it is hard to quantify the relative importance of what you can't measure relative to what you can," he says. "That's what we are facing with proteoforms." Perhaps proteoforms interact with one other, he says. This cooperative interaction may drive protein function, either on a per-molecule or on an ensemble basis.

The new single-molecule measurement platforms complement existing and emerging mass spectrometric methods, says Parag Mallick.

Making it easier to measure specific proteoforms will unlock important biology, he says. Much progress is happening related to measuring proteoforms, says Mallick, who highlights work in the Kelleher lab and others. Yet these approaches are challenging for most labs given the required experimental techniques, sophisticated instrumentation and complex data analysis, which add up to "barriers to entry for most labs."

Emerging single-molecule measurement platforms, like those at Nautilus Biotechnology, which he co-founded and where he is chief scientist, can measure, in a targeted manner, multiple proteoforms with single-molecule resolution.

In his view, these platforms complement existing and emerging mass spectrometric methods. In the longer term, he sees these platforms and the emerging nanopore sequencing approaches that operate on undigested proteins as ones that will support both targeted and discovery-style measurements. As the tools for broad-scale proteoform study improve, they will make their way into the broader research community. "But there remains significant methods development required to reduce the barrier to entry for these methods," says Mallick.

"A common thread is the need for more accessible tools," says Brian Reed.

Quantum-Si has rolled out its protein sequencing platform Platinum, and it's being used in a variety of ways, including protein engineering, peptide sequencing and antibody research, says Brian Reed who heads research at Quantum-Si. Some researchers add protein sequencing data to genomic and transcriptomic data they generate with high-throughput platforms. Overall, the platform is finding uses beyond traditional proteomics. "A common thread is the need for more accessible tools," he says, given that interrogating proteoforms with bottom-up and top-down mass spectrometry involve complex methods that take a high level of expertise, especially for detecting PTMs.

Platinum can detect PTM-based kinetic signatures. When the company's N-terminal amino acid recognizers bind to an amino acid that is modified or that is close to a PTM, the binding kinetics are influenced in a way that the sequencing chip can detect. For example, it can distinguish arginine and two types of arginine PTMs, dimethylation and citrullination, which can be challenging to detect with mass spec. The recent Second Annual Top-Down Proteomics Symposium led to a number of collaborations with Quantum-Si, he says, and it's motivating to him to take note of "the high level of interest among researchers in this space to adopt new methods like protein sequencing that can complement mass spectrometry."

An atlas and beyond

In 2021, Kelleher, Smith and colleagues in the Consortium for Top-Down Proteomics, which has around 500 members, presented the Human Proteoform Project[9] as "an ambitious initiative to define the human proteome; that is, to generate a definitive set of reference proteoforms produced from the genome." A letter of support signed by 56 scientists accompanies the paper. "We're really, really happy that these people endorsed this," says Paul Danis, the consortium's CEO. The letter is part of building community momentum. Findings matter, too, he says, such as work by Ying Ge at the University of Wisconsin and her group that shows how some proteoforms of cardiac proteins are more indicative of disease than others.

Given how complex human disease is, says Kelleher, "it is likely that proteoforms are the strongest connector between genotypes and phenotypes," a relationship modulated by other factors, too, such as metabolites and gene transcripts. To find the PTMs and isoforms that matter, "you have to know the proteoform landscape," as the Ge team explored with cardiac proteoforms. Compiling an atlas, says Kelleher, would give the community a "reference proteome." Without it, "you're just blind," and that slows down basic research, as well as drug development and ways to detect disease early.

Kelleher and his team have, for instance, mapped the proteoform landscape of the *KRAS* gene by using used top-down proteomics and immune enrichment to characterize 39 KRAS-derived proteoforms in colon cancer cell lines[10]. This proteoform landscape sheds new light on RAS biology, and the proteoform complexity gives hints for designing inhibitors of the KRAS protein. The lab has also published a proteoform atlas of five human tissues[11].

The word proteoform "would be a good addition to the biochemistry and biology dictionary," says Salvatore Sechi.

In the Smith lab, several proteoform-oriented projects are underway. One is a ribosome proteoform atlas and another is focused on HIV virion-related proteoforms, says Smith. The data analysis is daunting, but such projects are models to reveal how best to study the complex proteoform landscape.

As a postdoctoral fellow with Leroy Hood at the California Institute of Technology, Smith developed the first fluorescence-based DNA sequencer, which was commercialized by Applied Biosystems. When he gave talks about this work, he remembers considerable pushback. Labs typically used Sanger sequencing for gene analysis, which at the time involved incorporating radioactive nucleotides into a DNA sample and obtaining a gel electrophoresis-based readout. Many in his audience thought, says Smith, "I was trying to sell them on a large paperweight."

A few years later, there was no pushback since many were using the instrument, which automated classic Sanger sequencing and involved fluorescent dyes. "No one wanted to go back to radioactivity," he says. Proteomics could follow a similar trajectory. Sequencing human genes was "very laborious, very expensive" he says, and "boring as hell." New technologies have made sequencing genomes faster and easier.

"To me, proteoforms tell the whole story of a protein," says Benjamin Garcia.

When he saw that the NIH National Human Genome Research Institute (NHGRI) was looking for technologies that bridge the chasm between genotype and phenotype, says Smith, he thought: proteins are that bridge. Yet most projects in this space still involve delivering a sequencing readout. Among proteoform hunters, Smith hopes to see "a boiling cauldron" of new ideas, new efforts and new developments. "Just because mass spec is currently the best technology doesn't mean it has to be forever," he says. Mass spec certainly will be a piece of it, he says, "and then over time, maybe it'll phase out." For now, mass spec is the go-to technology for chasing proteoforms.

The scientists have been speaking with funding agencies and foundations, but at press time the Human Proteoform Project appears to lack firm funding commitments. NIH's Sechi says: "Briefly, for many reasons I preferred to not comment at this stage on the specifics of the proteoform atlas project."

It would be good, says Sechi, to have a more comprehensive characterization of the proteome, and this would imply a better characterization of the proteoforms of potential biomedical interest. "Definitely, we still have a lot of work to do before we can claim to have comprehensively characterized the human proteome," he says.

Even without a large-scale proteoform project in place, scientists continue their hunt for proteoforms. Success will take methods development and a bridge across the divide between subfields in proteomics. New approaches will play a role, too. Says Kelleher, the drop in the cost of genome sequencing and methods such as single-cell RNA sequencing have given labs a way to do biology-wide assays. "That is what proteomics lacks."

"I think it's needed," says Garcia about the atlas, and he signed the letter of support for the project. "It's time," almost 25 years after the Human Genome Project, to move to the next big challenge. Opinions differ on how to do this — for example, whether it should be a cell-based or disease-based project. "It will happen at some point, though," he says.

MacCoss likes the concept of a proteoform atlas. "I think it's a great idea," he says, and he, too, signed the letter of support. But the project's scale is challenging. For the human genome, the scientific community had a good estimate of size and scale, but with a project devoted to mapping all proteoforms, "we don't know what the endpoint is." He draws a comparison to the 'war on cancer' declared during the administration of US President Nixon and President Kennedy's plan for lunar space missions. For the former, much was unknown — certainly little was known about cancer's molecular underpinnings; the other had tangible, albeit daunting, dimensions. "We knew where the moon is; we knew the engineering challenges to go to the moon."

With proteoforms, the challenge he sees is that a protein can exist in so many variants, with any number and combinations of post-translational modifications.

If the overall cost of doing proteomics were dramatically lower, says MIchael MacCoss, "that would then open up the game a little bit more."

Perhaps, he says, a smaller-scale, technology-focused project such as the $1,000 Genome Program has advantages for the proteoform hunt. This National Human Genome Research Institute program launched in 2004 was focused on technology development. It aimed to "dramatically reduce the cost of DNA sequencing." Cheaper technology will make it easier to hunt proteoforms. Then more labs could join in, he says, and collaborations with companies could emerge, too. If the overall cost of doing proteomics were dramatically lower, he says, "that would then open up the game a little bit more."

Vivien Marx ✉
Nature Methods.
✉e-mail: v.marx@us.nature.com

Published online: 2 February 2024

References

1. Po, A. & Eyers, C. E. *J. Proteome Res.* **22**, 3663–3675 (2023).
2. Aebersold, R. et al. *Nat. Chem. Biol.* **14**, 206–214 (2018).
3. Smith, L. M., Kelleher, N. L. & Consortium for Top Down Proteomics. *Nat. Methods* **10**, 186–187 (2013).
4. Kelleher, N. L. et al. *J. Am. Chem Soc.* **121**, 806–812 (1999).
5. Eng, J. K., McCormack, A. L. & Yates, J. R. III *J. Am. Soc. Mass Spectrom.* **5**, 976–989 (1994).
6. Plubell, D. L. et al. *J. Proteome Res.* **21**, 891–898 (2022).
7. Lutomski, C. A. *Angew. Chem. Int. Ed. Engl.* **62**, e20230594 (2023).
8. Chen, S. et al. *Nature* **604**, 384–390 (2022).
9. Smith, L. M. et al. *Sci. Adv.* **7**, eabk0734 (2021).
10. Adams, L. M. et al. *J. Biol. Chem.* **299**, 102768 (2023).
11. Drown, B. S. et al. *J. Proteome Res.* **21**, 1299–1310 (2022).