

# Proteomics: taking on protein complexes

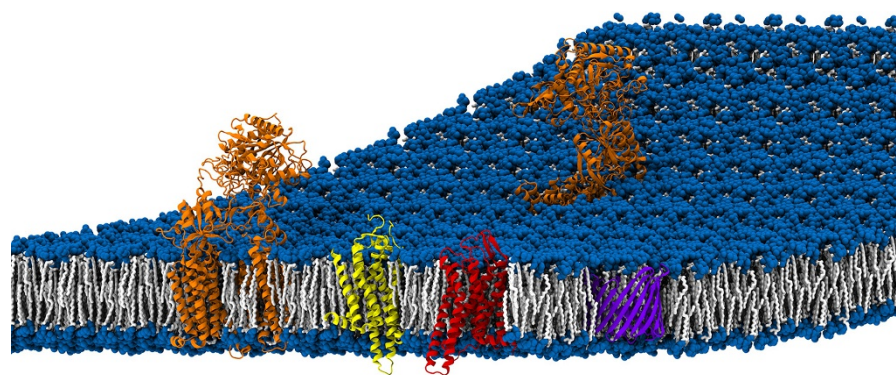
Vivien Marx

With mass spectrometry and grit, labs discover structural and functional secrets about protein complexes.

Proteins collaborate. They hook up into protein complexes with few or many collaborators in liaisons that are brief, longish or long-term. To understand what these protein complexes do, labs seek high-resolution structural views. Giant protein complexes are among the most difficult ones to study, says Carol Robinson of the University of Oxford, but very small membrane complexes are “notoriously unstable,” which makes drug-binding experiments challenging, as they often involve membrane complexes.

No single technique is likely to help labs understand how evolution shaped a given protein complex and its function, says Rockefeller University researcher Brian Chait. But a single tool could one day help to characterize the structure of a protein complex on an atomic scale: an X-ray microscope. His colleague and long-time collaborator Michael Rout dreams of using such an instrument to create an interactive multi-scale map in which a researcher can move visually around an atomic-scale rendering of a protein complex and on a time scale from nanoseconds to a second. Although it is possible to achieve such a map with currently available methods, “it’s also easy to get it shockingly wrong,” Rout says. One could dream further still and consider a map of this type of the entire cell, says Chait. “I dream about it, anyway.”

Until labs have such an instrument delivering atomic-level data about any protein complex of interest, they must use multiple approaches, each with fortes and caveats. X-ray crystallography and nuclear magnetic resonance (NMR) help scientists obtain three-dimensional structural information. But these methods require much material, and they typically involve recombinant protein overexpression, which can skew the view of a protein complex in a



M. Degiacomi, Robinson lab, University of Oxford

A cell has many types of membrane protein complexes, which were long not amenable to mass spectrometry. That is slowly changing. Native mass spectrometry can help to show a small molecule binding to membrane proteins, says Carol Robinson.

living cell. With electron microscopy it can be hard to detect the binding of a small molecule to a large protein assembly, says Robinson. Confocal microscopy shows protein movement, but not on an atomic scale.

Mass spectrometry delivers what is hard to get in other ways: the chemical composition of a protein complex, says Northwestern University researcher Neil Kelleher. The approach uses little sample and delivers high-resolution data without the need to perform protein overexpression, says Albert Heck of the University of Utrecht. Labs can do native mass spectrometry in which the complex is gingerly removed from the cell and kept intact, or, with cross-linking mass spectrometry, they can first lock specific protein residues in place<sup>1-3</sup>.

“I think the strength of all of these approaches is often in their combination,” says Robinson. Cross-linking mass spec is powerful for defining protein-interaction sites, she says. To get stoichiometry data, which is the count of protein subunits in a given complex, or for looking at a small-

molecule binding to membrane proteins, “native mass spectrometry has tremendous advantages in terms of amount of protein required, information obtained: measurement of dissociation constants, competitive binding experiments, distinguishing lipid from drug binding—to name but a few,” she says.

## Going native

Ideally, researchers want to watch the protein complex “in its native environment just doing its thing,” says Chait. Mass spec helps labs do so with instruments that are either bought or built in the lab. For measuring protein complexes and protein stoichiometry, the “unequivocal power of native mass spectrometry” comes into play, he says. As long as the measurement is not distorted and there is good fidelity to the native protein complex, “you’re getting an answer that you can perhaps believe in.”

That fidelity to the native protein complex is hard to achieve. Chait, his Rockefeller colleague Rout and others applied immuno-electron microscopy, affinity purification, mass spec and

computational methods to characterize the yeast nuclear pore complex (NPC) and found that it has 456 protein molecules and 30 distinct proteins<sup>4</sup>. They know much about the NPC: its mass is around 20 times that of a ribosome; it has spokes and symmetry; it has, as Rout says, an “ancient fossil” at its core, which makes it a cousin of vesicle-coating protein complexes that predate the eukaryotic cell; and it stabilizes curved membranes. But how the NPC handles transport in and out of the nucleus is still under scrutiny, and a lot of structural knowledge is still, says Rout, the equivalent of a medieval map with notations such as “here be dragons.”

It's tricky to keep a protein complex intact. Robinson and her group have used native mass spectrometry to generate high-resolution data for membrane complexes. For example, they modified the Orbitrap mass spectrometry platform to be able to distinguish lipid, peptide and drug binding to a membrane protein complex. They had to find the best way to release the protein complex from micelles—a detergent assembly that both stabilizes the protein complex after its removal from the cell and keeps the complex relatively intact in the mass spectrometer. Buffers can be challenging: what is needed to remove the complex from the cell is not what the instrument needs. “By all means take the complex out in a particular buffer and we will change it for one that we like better,” says Robinson. “We can usually do this without losing the complex.” It takes trial and error to explore solutions with varying ionic strength or pH, she says. Techniques and new mass-spec-compatible buffers and detergents keep emerging.

Up until a few years ago, Robinson says, she would not have been able to help a structural biologist colleague seeking to study the binding of a ligand to an understudied membrane complex that does not crystallize. “But nowadays I would be much more positive,” she says. The protocol to help this colleague would include looking at the behavior of the membrane protein in the mass spectrometer, detergent screening, adjusting conditions, exploring how the protein responds to ligands and taking some specific drug-binding measurements. A lab will also want to see whether the mass spec data agree with molecular dynamics simulations. “It is quite an iterative process and



Northwestern University

Native mass spec using intact protein complexes will bring better understanding of biological variation in protein complexes, says Neil Kelleher.

often we can't predict how something will behave in a particular detergent or lipid background,” she says. “Maybe one day we will be able to predict this, but for now it is very much a ‘suck it and see’ process.”

In the mass spectrometer, the sample is ionized, releasing the pro-

tein complex from the aqueous environment into the gas phase. After proteins or protein complexes are broken into fragments, their patterns of mass-to-charge ratios are used to identify them. In tweaking experimental parameters, Robinson and her group have found, for example, that a combination of optimized pressures and a gentle voltage gradient help keep proteins in their folded state as they leave solution and reach the gas phase. Years ago she once took a drill to her mass spec instrument to increase pressure in the early stages of the flight path to create plenty of soft collisions with gas molecules, dampening the energy of the complexes so they didn't fall apart. Such drastic measures are no longer needed today, but, she advises labs, “don't be afraid to experiment with conditions.”

A great advantage of native mass spec, says Robinson, is that it makes it possible to study proteins within the context of larger assemblies. Given how heterogeneous assemblies are, they are challenging to study crystallographically. Native mass spec can measure picomole amounts of a protein complex that has retained its quaternary structure, and it offers a way to measure protein complexes with a mixture of different types of molecules.

Chait praises the work done in the many labs using native mass spec, but his sense is that the technique “has underperformed terribly for endogenous complexes.” The stoichiometric data are not as good as needed, and experiments fail with many complexes. Ongoing research will still clarify plenty, but he is concerned when insufficient data invite the community to guess too much about protein complexes.

Heck coined the term ‘native mass spec’ and has long fielded questions about

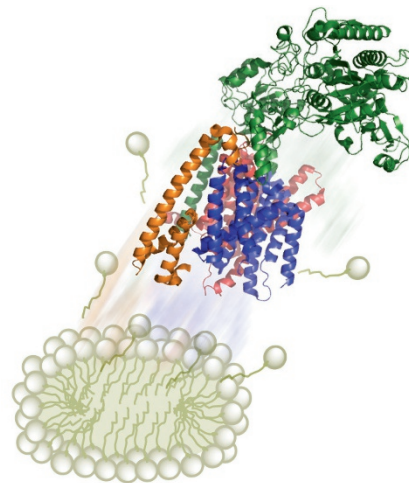
whether it implies *in vivo* measurement of protein complexes. He invariably responds that was never the intention. Rather, he and his team sought to describe an assay that kept proteins intact, and they appropriated the term from protein separation run on ‘native’ gels.

‘Native’ is taking on a new dimension: the heterogeneity of protein complexes. As Rout explains, protein complexes are likely to change with a cell's metabolic state. A cell might have various NPCs, and NPCs might differ between cell types. As Kelleher explains, this diversity of proteoform complexes has an informal moniker: ‘complex-o-forms’.

### Cross-linking mass spec

The palpable excitement about cross-linking mass spectrometry (XL-MS) is warranted, says Chait. The principle is old, but there have been technological limitations. XL-MS requires high-dynamic-range instruments. With XL-MS, chemical linkers connect residues in a protein complex. The structural information obtained is based on the peptides identified using the cleaved linkers as reporters. The cross-linker sites yield the distance between two residues. Labs can explore subunits within the folded complex, study the complex's topology and see how structure changes under different conditions. Cross-linkers added to a sample distribute randomly and then link residues. “The trouble is, it's incredibly inefficient,” says Chait.

“I love to tackle complexes that have proved intractable using other methods,”



It's tricky to release a membrane protein complex from a detergent assembly into the gas phase to obtain a precise mass measurement and track its conformational changes.

J. Hopper, Robinson lab, University of Oxford; Erin Dewalt/Nature Publishing Group



M. Riffle, Davis lab, University of Washington

Using software called ProXL, researchers visualize cross-links in a yeast protein complex.

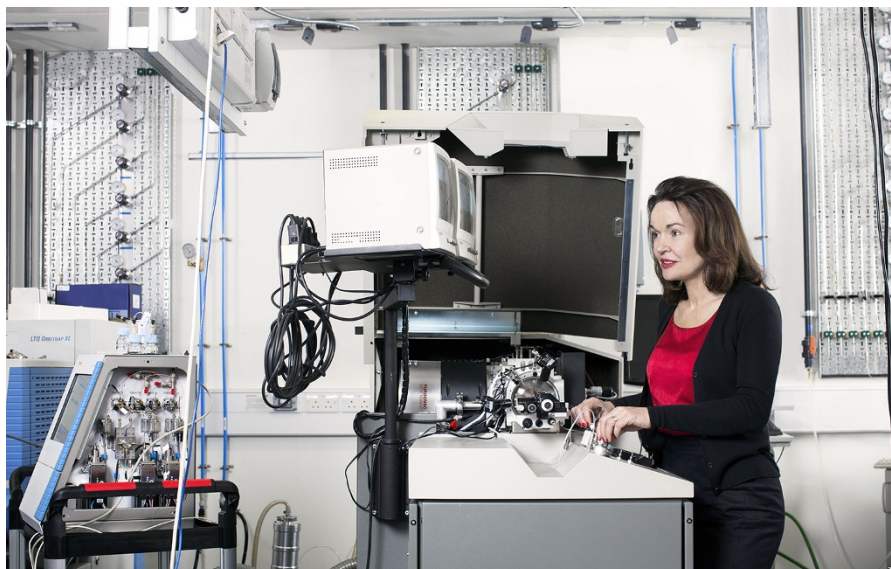
says Alex Zelter, a biochemist in Trisha Davis's lab at the University of Washington and who uses XL-MS. Cross-linking data can add missing puzzle pieces to complete a structural picture, such as when a lab has a low-resolution, 20–30-angstrom electron microscopy structure or perhaps crystal or NMR structures of individual domains in a complex. XL-MS data can be used to computationally generate a structure without additional data and, for example, help labs assess protein therapeutics or explore basic research questions. "XL-MS can provide a great deal of structural information in just a few hours of MS time, allowing rapid assessment of protein structure and conformation," he says.

Just a few micrograms of unpurified protein can deliver useful data from XL-MS, says Zelter. "One of the beauties of XL-MS is that it freezes the structural information prior to analysis," he says. Theoretically, the complex can even be cross-linked in the living cell. After cross-linking, researchers can be "unfriendly" to the target protein complex: they can strip away mass-spec-incompatible reagents or enrich proteins or peptides of interest. "No worries," he says; the cross-links were formed in the "benign" environment,

so these steps won't affect the structural data obtained later. XL-MS works well on large complexes, says Zelter, and it can trap interactions in transient complexes. Integrative modeling software is helping to advance the way XL-MS data can be added to other forms of structural data.

Enrichment of protein complexes is an art, says Michael MacCoss of the University of Washington, who collaborates with Zelter. "What might work for one might not necessarily work for another," he says. When looking at protein complexes, says MacCoss, experimental controls are key: positive controls make sure the experimental steps worked as expected, and negative controls help to distinguish between abundant proteins and those co-enriched as an artifact of the sample prep and the *bona fide* members of the complex. Better quantitative mass spectrometry is making these steps more routine.

As Trisha Davis explains, what she likes best about XL-MS is that it can provide structural information with relatively little protein, which doesn't even have to be that pure. A few years ago XL-MS experiments led to a mere handful of cross-links from a given protein complex, and, says Zelter,



Brigitte Lacombe

Patience and determination are needed to get mass spectra of difficult protein complexes, says Carol Robinson.

“we are now routinely identifying hundreds of links.” When structural biologists approach with challenging problems, “our answer has therefore changed from something like, ‘We can try, and perhaps we’ll identify a key cross-link that will help you’ to ‘Yes! We can get a lot of structural information for you about your complex and we can do it in short order.’”

MacCoss says considerable effort is required to get this process working well given that “you want the complex covalently linked but not in a huge aggregated mess.” Now that mass spec instruments have higher resolution, XL-MS data quality has improved, says MacCoss. Kojak, an algorithm developed in collaboration with colleagues by Michael Hoopmann, who is in Robert Moritz’s lab at the Institute for Systems Biology, has also convinced MacCoss that it is possible to automate interpretation of cross-linked peptide tandem mass spectrometry spectra<sup>5</sup>.

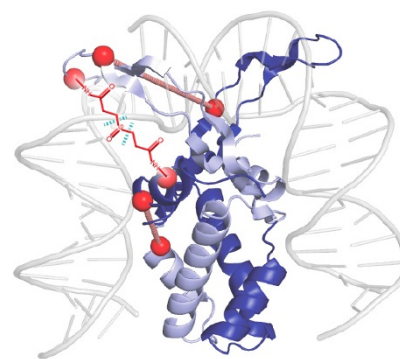
### Going proteome-wide

A number of labs take a ‘proteome-wide’ approach to cross-linking mass spec experiments on protein complexes. Heck and his team recently detected over 2,000 cross-linked peptides from lysed HeLa cells and investigated intraprotein and interprotein links, leading to topological information about the protein assemblies<sup>6</sup>.

The scientists screened the mass spec readout by searching a human proteome database using a dedicated tool called XlinkX. They found hundreds of

thousands of peptides and whittled down the search to find the true cross-linked peptides, sorting out the artifacts and false positive ‘hits’. Heck is glad about the number of detected cross-linked peptides, but, he says, the experimental outcome harkens back to proteomics a decade ago when labs were happy when they detected 100 proteins in the cell. Over 90% of the detected proteins were among the 500 most abundant proteins in HeLa cells, he says. “I actually think that underneath our data there are thousands more cross-links,” he says. The data are “just a start.”

Juri Rappsilber of the Wellcome Trust Center for Cell Biology at the University of Edinburgh points to XL-MS positives and challenges<sup>7</sup>. The technique can be “undemocratic,” when it yields more data



F. Liu, Heck Lab, University of Utrecht; Erin Dewalt/Nature Publishing Group

With cross-linking mass spec, linkers connect residues in a protein complex and chemically tie them together. Structural information is based on the identified cross-linked peptides; the cleaved linkers act as reporters.

for some parts of a protein or a protein complex than others. Cross-linking can trigger physiological change, and artifacts can occur when a protein is captured in a rare conformation.

Yes, says Zelter, some protein residues may cross-link more readily, but it remains true that any two cross-linked residues came within cross-linking distance. Still, the complex might have ‘dark regions’ that lack linkable residues; some peptides do not perform well during chromatography or in the mass spectrometer. “This does not bias your view so long as you do not draw conclusions from the absence of observations,” he says. Researchers can use multiple cross-linkers with different reaction chemistries to help address any bias issues. They can see whether a residue was reactive by looking for monolinks, which are cross-linkers that reacted only with a single residue.

XL-MS brings on ‘big data’ challenges: in an experiment with 100,000 peptides, the 500 cross-linked ones need to be found, says Heck. “That’s very hard because they are not so different from the other peptides,” he says. One algorithmic approach he is exploring is to reduce the search space for cross-linked peptides. Rather than have an algorithm keep searching for a specific fragmentation pattern indicative of a cross-linked peptide, the mass spectrometer essentially rests until the specific marker ion is detected, and only then is a peptide sequenced to generate data about cross-linked peptides, says Heck. He is currently in negotiations with vendors about commercializing one of his algorithms. Another project of his is to enable viewing of the cross-links on Protein Data Bank structures.

### Weighing options

Typically, a linker reaches out to residues with two arms. There are also three-armed cross-linkers to which an affinity tag can be attached with which to pull down cross-linked peptides. These three-armed linkers seem promising, says Heck, but in experiments they are not yet performing better than the two-armed ones. “I think there is still a fair amount of lab lore involved in this,” says MacCoss. Some labs swear by custom cross-linking reagents with affinity tags, others insist on using a variety of cross-linkers. The choice depends on the research question, he says—for example, whether a scientist



University of Utrecht

Native mass spec, XL-MS and electron microscopy are an ideal threesome, says Albert Heck.

is looking for proteins that interact or is interested in providing a structural constraint, or whether cross-linking happens on an enriched complex or an unfractionated lysate.

Chait says that XL-MS is useful but still early-stage, and the limitations

indeed include protein “hot spots” that preferentially cross-link. Better algorithms will help labs sort and integrate data. In their experiments, says Rout, they found that cross-linkers missed sets of expected interactions.

In Robinson’s view, cross-linking “in theory is very straightforward but in practice requires careful control of concentrations of proteins and cross-linkers.” She also highlights difficulties in data analysis that make it a challenging research area. “Too much cross-linker—everything is linked,” she says. Distinguishing cross-linking within one protein and cross-linking between two copies of the same protein subunit can also be difficult. The computational algorithms continue to improve, she says. The approaches will benefit from more method integration, bringing together cross-linking and native mass spec to assist with interpretation of electron microscopy data.

XL-MS and cross-linker chemistry will keep advancing, says Kelleher, and providing more information about the folds of recombinant and endogenous protein assemblies. In his view, a better understanding of biological variation in protein complexes will come from native mass spec using intact protein complexes. XL-MS takes a common route in proteomics by digesting proteins, then analyzing peptides and inferring structure. Believing that labs can get to complete compositional information by measuring enough peptides “is a flawed world view,” he says. Peptide identification permits inference about structure and genes, he says, but a native ‘top-down’ mass spec approach that avoids protein digestion and looks at things intact would also avoid inference and deliver greater molecular specificity about a protein complex, for example, of all the subunits and their



Regine Schulte

When setting up mass spectrometers to analyze protein complexes, challenges include achieving sensitivity and resolution for high masses, says Uwe Schulte.

proteoforms. In that sense it could deliver information on gene products and enable labs to better study biological variation.

Kelleher and his team recently presented a native top-down proteomics informatics approach to enable integrated analysis of data from intact protein

complexes<sup>8</sup>. In his view, it's the beginning of doing proteomics "completely native," he says, while preserving the co-factors, protein-protein interactions and post-translational modifications during mass spec analysis, "all of these things that are closer to function and closer to the endogenous sample." Routine analysis of protein complexes in this way could be possible in 3–5 years, he says. Kelleher says the platform will still undergo much refinement; he and his team are exploring how to use it to determine stoichiometry and modifications. The approach was developed in collaboration with Thermo Fisher Scientific, and the company commercialized parts of the software earlier this year.

### Drug discovery headaches

Robinson has long been collaborating with pharmaceutical and biotechnology companies and has just formalized those relationships by setting up a company in Oxford called OMass Technologies. The idea is to help drug developers tackle difficult projects with approaches from her lab. The OMass core scientific team consists of three members of her lab who have much experience working together, and she hopes the company can grow and expand.

Logopharm in March-Buchheim, Germany, uses mass spec to help pharma and biotech researchers with projects related to membrane proteins and protein complexes in a context of pressures from growing R&D costs and prominent drug failures. Complexity at the molecular level challenges labs seeking a mechanistic understanding of a protein complex, says Uwe Schulte, the company's managing director. Academics approach the company for exploratory work; industry customers do so

with some ideas about the mechanism or function of a target protein and often with the goal of identifying changes or quantities of known compounds or markers, targets and pathways under a variety of conditions.

Co-crystals are used heavily in drug design, but when it comes to protein complexes, crystal structures usually are not available or do not reflect conformational dynamics well, says Schulte. At Logopharm, the approach is to solubilize proteins with detergent-containing buffers and then perform affinity purification followed by liquid chromatography-tandem mass spectrometry. A delicate balance must be struck: some proteins resist being solubilized, and a single buffer might not equally preserve all interactions. A membrane-disrupting detergent can disrupt very weak or dynamic interactions. The company scientists developed a proprietary buffer system with an eye to these issues as well as protein complex integrity, he says.

When setting up mass spectrometers, some of the technical challenges to address are achieving sensitivity and resolution for high masses and issues related to the limited stability of large noncovalent complexes in the ion-conducting pathway and detector.

When taking a top-down approach to protein complexes, labs face plenty of biochemical difficulties, such as inadequate solubility, stability, purity, homogeneity and abundance of protein complexes, says Schulte. To date only a few labs have been able to characterize particularly stable membrane protein complexes. His experience with cross-linkers has been "rather negative," and one weakness of cross-linkers is that they probe proximity but not biochemical interaction; they can work well with some proteins and fail miserably with others, which leads to a high false negative rate. There are issues with cross-linker concentrations that lead to too much unspecific cross-linking. XL-MS has potential as a way to deliver structural information about protein complexes. "But that is something very different from unbiased identification of protein complexes in their native environment," he says.

### All techniques on deck

Native mass spec, XL-MS and electron microscopy are an ideal threesome, says Heck, given that they can handle the same types of biological structures, the

same sample amounts. As these techniques nudge closer, skill requirements for proteomics researchers are broadening. In the past, a researcher could be a good mass spectrometrist or a good crystallographer or NMR spectroscopist, but the needs of modern structural biology mean that scientists must draw on a number of techniques. Heck and his team collaborate with electron microscopists, for example, and he says he often hears the adage ‘the future in structural biology is hybrid’.

The hybrid future is a challenge for scientists early in their careers, says Heck. Were he just starting out, he would do as his students do, which is to establish a background in mass spec and then, for example, do a postdoctoral fellowship in electron microscopy. That combination, he says, is “almost a ticket for a new job.” Some experience will be helpful, even if it is not high-level expertise in electron microscopy or XL-MS.

As protein complex structure is explored, the emerging picture is dynamic, not static, and structure characterizations need to reflect this, says Heck. The ribosome, for example, is made up of a stable core of proteins and others that come and go, much like a factory with a busy loading dock.

Chait and Rout want atomic detail about the NPC structure and to learn how it changes with the cargo it handles, how components leave or join the complex, how the NPC might be rebuilt in the lab. They use many techniques; lately they have been trying NMR to learn about the NPC’s many disordered proteins that do not crystallize well. “You use whatever you can,” says Chait.

It certainly takes much patience and determination to get mass spectra of difficult protein complexes, says Robinson. Most researchers in her group are skilled at protein expression, including of proteins from membranes, as well as at running the

mass spectrometers and molecular modeling. After lectures she often hears from members of the audience that she has made the work look too easy because in reality it is hard to get these experiments to work. Her response: “I agree with this but I don’t want to spend my lectures highlighting the problems—just the exciting new things that can be done.”

1. Marcoux, J. & Robinson, C.V. *Structure* **21**, 1541–1550 (2013).
2. Aebersold, R. & Mann, M. *Nature* **422**, 198–207 (2003).
3. Chait, B. *et al. J. Am. Soc. Mass Spectrom.* **27**, 952–965 (2016).
4. Alber, F. *et al. Nature* **450**, 695–701 (2007).
5. Hoopmann, M.R. *et al. J. Proteome Res.* **14**, 2190–2198 (2015).
6. Liu, F. *et al. Nat. Methods* **12**, 1179–1184 (2015).
7. Rappsilber, J. *J. Struct. Biol.* **173**, 530–540 (2011).
8. Skinner, O.S. *et al. Nat. Methods* **13**, 237–240 (2016).

---

**Vivien Marx is technology editor for *Nature Methods* ([v.marx@us.nature.com](mailto:v.marx@us.nature.com)).**