



Piecing together a structural puzzle

Sample preparation and instrumentation for macromolecular mass spectrometry continue to evolve, making the technique more relevant than ever.

“ It is often the case that a challenging scientific problem is not addressable directly but is instead solvable only by elegant inference ”

Mass spectrometry was first developed with the simple goal of measuring the mass — or rather the mass-to-charge ratio — of monoatomic ions¹. Although rudimentary, the first mass spectrograph could resolve a beam of neon ions into neon-20 and neon-22, a finding that led to the discovery of stable isotopes, no less^{2,3}. The following century would see improvements in ionization, analysis and detection that today make mass spectrometry applicable to a wide range of analytes, including molecules that are very large and fragile. These improvements include notable work by John Fenn and colleagues, who developed “electrospray wings for molecular elephants” to gently charge and volatilize proteins so that they could fly to a detector intact⁴.

Mass spectrometry has become an indispensable tool in proteomics not only to measure intact polypeptide ions in the gas phase but also to trap and fragment them residue-by-residue, thereby enabling us to identify each constituent and indirectly unravel the original amino acid sequence. This fragmentation method, known as tandem mass spectrometry, is a nice example of how an elaborate molecule can often be converted (chemically or otherwise) into more tractable secondary species that inform us about the original molecule. The strategy can be extended to glycoproteins, which have the complexity of a protein with the added diversity of pendant oligosaccharide chains and their many possible linkage isomers. A Review by Noortje de Haan and co-workers⁵ describes how the difficulties in elucidating the structure of sialic acid glycoproteins can be overcome by subjecting them to functionalization reactions in which the outcomes depend on how the glycoproteins are linked together. The nature of the resulting products thus informs us about the parent glycoproteins, thereby giving us a well-earned piece of a complex structural puzzle.

As with biomolecules, synthetic polymers are also often amenable to mass spectrometry, the data from which complement solution techniques such as gel permeation chromatography, which can tell us about molar mass distributions⁴. Aside from electrospray ionization (ESI), matrix-assisted laser desorption ionization (MALDI) is the source of choice in the synthetic polymer realm, this being despite the pioneering work of Koichi Tanaka and colleagues having first targeted proteins⁶. The sources and conditions his team developed, relative to ESI, afford ions with lower charge, such that they often appear as a single ‘envelope’ in a

mass spectrum. However, a Perspective by Christopher Barner-Kowollik and co-authors⁷ highlights a substantial weakness inherent to MALDI — it is not readily compatible with flow solution inlets and this limits its applications in online monitoring. Moreover, there is the eternal question of whether the speciation of macromolecules in the gas phase reflects that in the solution phase. Regardless of the ionization source, different molecules just have different ‘flyabilities’⁸, which is to say that some elephants are more likely to take to the skies than others. This touchy subject is undoubtedly worthy of attention because the exact mechanisms by which analytes undergo ionization are not fully understood.

Many of us are aware of the sensitivity of mass spectrometry, but we should not be fooled into thinking that present instrumentation is anywhere near optimal. For example, of the analyte molecules that undergo electrospray ionization, even at very low flow rates (‘nanoelectrospray’) only one in a few hundred reaches the detector⁹. This fraction drops precipitously to one in hundreds of thousands when the analyte solution is introduced at higher flow rates. Along with other practical limitations, including resolving power and mass-to-charge range, it is clear that there is work to be done. Some of this optimization does not concern the mass spectrometer itself but rather what goes in it⁷. Indeed, extracting quantitative data from mixtures of analytes usually calls for separation before ionization, with various chromatographic techniques being at our disposal. It is also possible to separate analytes after ionization but before analysis. For example, ion-mobility spectrometry is now a mainstream way to separate macromolecular ions according to the times they take to move through an inert gas. This method, which is essentially electrophoresis in the gas phase, is conceptually simple yet practically challenging. For example, measuring the drift time of an ion can enable one to compute its collision cross-section, but this is typically a relative rather than an absolute size. We do not get a direct picture of the ion nor do we even know if its conformation or size in the gas phase reflects the native structure. What we get is another piece in the puzzle, some indirect information from which we can say something — no more or less — about our macromolecules.

‘Indirect characterization’, in which one exploits the chemical properties of candidate structures to deduce a compound’s identity, is an intellectual pursuit that

harkens back to the days when chemists had new compounds in their hands but few analytical instruments in their labs. A shrewd chemist might have pieced together a structural puzzle by considering the physical or chemical behaviour of a compound. Perhaps they inferred the nature of an alcohol group by testing its propensity to undergo oxidation (or lactonization, in the case of glycoproteins). Nowadays, at least for simple molecules, one would forego chemical tests, collect a NMR spectrum and declare victory. A complex molecule, however, presents a characterization puzzle that will take longer to solve and might need to be approached from different angles. It is often the case that a challenging scientific problem is not addressable directly but is instead solvable only by elegant inference. For example, we are unable, no matter how well we train our eyes on the night sky, to spy planets beyond our solar system. But we know they exist because each exoplanet exerts a gravitational pull on its neighbouring star, causing a Doppler shift we definitely can observe.

We live in a time in which we can ‘see’ single molecules. Characterization just does not get any more direct than seeing a molecule with a scanning probe microscope

or transmission electron microscope. Although there is undeniable pleasure in gazing at such micrographs, just as much satisfaction can be had when discovering that the same molecule — or even part of it — gives rise to a peak in a spectrum. It is good to know that the applications of mass spectrometry are more numerous than ever and that the method retains relevance and independent value that will persist as long as our curiosity drives us to solve scientific puzzles.

1. Doerr, A. et al. Editorial. *Nat. Methods* **12**, 3 (2015).
2. Thomson, J. J. Bakerian lecture: rays of positive electricity. *Proc. Roy. Soc. A* **89**, 1–20 (1913).
3. Aston, F. W. A positive ray spectrograph. *Philos. Mag.* **38**, 707–714 (1919).
4. Fenn, J. B. Electrospray wings for molecular elephants (Nobel lecture). *Angew. Chem. Int. Ed.* **42**, 3871–3894 (2003).
5. de Haan, N. et al. Glycomics studies using sialic acid derivatization and mass spectrometry. *Nat. Rev. Chem.* <https://doi.org/10.1038/s41570-020-0174-3> (2020).
6. Tanaka, K. The origin of macromolecule ionization by laser irradiation (Nobel lecture). *Angew. Chem. Int. Ed.* **42**, 3860–3870 (2003).
7. De Bruycker, K. et al. Mass spectrometry as a tool to advance polymer science. *Nat. Rev. Chem.* <https://doi.org/10.1038/s41570-020-0168-1> (2020).
8. Mann, M. The ever expanding scope of electrospray mass spectrometry — a 30 year journey. *Nat. Commun.* **10**, 3744 (2019).