MILESTONE 9

Solving the primary structure of peptides

By the late 1950s, chemists had realized that mass spectrometry could be used to decipher the structures of molecules (**Milestone 7**). Scientists were also just beginning to identify the primary structure, or sequence, of peptides and proteins using chemical approaches. In 1953, Fred Sanger used N-terminal labeling of peptide fragments, followed by hydrolysis and analysis via paper chromatography, to sequence insulin (a feat that earned him the Nobel Prize in Chemistry in 1958). Around the same time, Pehr Edman devised a method for sequencing proteins by stepwise degradation starting from their N termini.

The application of mass spectrometry to peptide sequencing came shortly thereafter, in 1959, when Klaus Biemann and colleagues described an innovative way to elucidate peptide structures using the reduction of small peptides to generate polyamino alcohols with characteristic spectra. Their key advance was identifying chemistry to reduce the highly polar, zwitterionic character of peptides to allow them to be vaporized for ionization. In the years that followed, numerous groups pioneered methods for mass spectrometry–based peptide sequencing, developing ways to chemically modify peptides to be compatible with technical innovations that allowed direct introduction of samples into the ion source for mass spectrometry.

In parallel, Biemann and colleagues built on their previous work to develop a sequencing strategy that was both fast and generally applicable for use on short peptides. However, they soon found that the mass spectra were complicated by factors such as side-chain fragmentation and variable ion abundance. To sort these out, they used a computational approach to interpret the mass spectra. The technique relied on using the exact masses of ion fragments to compute all possible peptide sequences, from which they could select the most probable sequence on the basis of the most abundant ions. They confirmed their choice by looking for other ions that should be present were the selected structure correct. This work was among the first to use computers to analyze mass spectra, setting an important precedent for the field.

The late 1960s and 1970s saw many advances in peptide sequencing by mass spectrometry, including a permethylation technique that Howard Morris and colleagues used for partial sequencing of proteins. However, mass spectrometry methods had competition from alternative approaches, such as Edman sequencing—which was streamlined by that point—as well as DNA sequencing, which yielded gene sequences that could be translated into the protein sequence. Ultimately, mass spectrometry proved complementary to these techniques, allowing researchers to determine the C termini of proteins that were too long for Edman sequencing and to confirm translated sequences.

The early 1980s brought further innovations. In 1981, Donald Hunt and co-workers carried out the first sequencing of peptides by tandem mass spectrometry (**Milestone 13**). They analyzed permethylated peptides on a triple quadrupole (**Milestone 6**) mass spectrometer following chemical ionization; this allowed direct analysis of a complex mixture of peptides, generated by protease cleavage of a large protein, without prior fractionation. Tandem mass spectrometry soon became the standard method for peptide sequencing.

Another revolution was the introduction of 'soft' ionization methods, which can be used on polar, thermally labile compounds and yield ions that are not highly fragmented (see **Milestone 2**). In 1981, Michael Barber and colleagues developed one of the first of these techniques: fast atom bombardment (FAB), which involves mixing samples in solution with a matrix and bombarding them with high-energy atoms. FAB allowed the group to sequence unmodified peptides. Although important, FAB was ultimately surpassed by soft ionization methods such as matrix-assisted laser-desorption/ionization (MALDI; **Milestone 18**), which are used widely today.

The past few decades have been fruitful, and the use of mass spectrometry for peptide and protein analysis has become commonplace. It is clear that these and other seminal works have had a lasting impact on the analysis of protein primary structures.

Rita Strack, Assistant Editor, Nature Methods

ORIGINAL RESEARCH PAPERS Biemann, K., Gapp, G. & Seibl, J. Application of mass spectrometry to structure problems. I. Amino acid sequence in peptides. *J. Am. Chem. Soc.* **81**, 2274–2275 (1959) | Biemann, K., Cone, C., Webster, B.R. & Arsenault, G.P. Determination of the amino acid sequence in oligopeptides by computer interpretation of their high-resolution mass spectra. *J. Am. Chem. Soc.* **88**, 5598–1606 (1966) | Hunt, D.F., Buko, A.M., Ballard, J.M., Shabanowitz, J. & Giordani, A.B. Sequence analysis of polypeptides by collision activated dissociation on a triple quadrupole mass spectrometer. *Biomed. Mass Spectrom.* **8**, 397–408 (1981) | Barber, M., Bordoli, R.S., Sedgwick, R.D. & Tyler, A.N. Fast atom bombardment of solids as an ion source in mass spectrometry. *Nature* **293**, 270–275 (1981)
FURTHER READING Morris, H.R., Williams, D.H. & Ambler, R.P. Determination of the sequences of protein-derived peptides and peptide mixtures by mass spectrometry. *Biochem J.* **125**, 189–201 (1971)