

Technical tips

Session 9

Mass spectrometry for protein identification: In mass spectrometry, a substance is bombarded with an electron beam having sufficient energy to fragment the molecule. The positive fragments which are produced (cations and radical cations) are accelerated in a vacuum through a magnetic field and are sorted on the basis of mass-to-charge ratio. Since the bulk of the ions produced in the mass spectrometer carry a unit positive charge, the value m/e is equivalent to the molecular weight of the fragment. The analysis of mass spectroscopy information involves the re-assembling of fragments, working backwards to generate the original molecule. A schematic representation of a mass spectrometer is shown below:

(Image removed due to copyright reasons.)

A very low concentration of sample molecules is allowed to leak into the ionization chamber (which is under a very high vacuum) where they are bombarded by a high-energy electron beam. The molecules fragment and the positive ions produced are accelerated through a charged array into an analyzing tube. The path of the charged molecules is bent by an applied magnetic field. Ions having low mass (low momentum) will be deflected most by this field and will collide with the walls of the analyzer. Likewise, high momentum ions will not be deflected enough and will also collide with the analyzer wall. Ions having the proper mass-to-charge ratio, however, will follow the path of the analyzer, exit through the slit and collide with the Collector. This generates an electric current, which is then amplified and detected. By varying the strength of the magnetic field, the mass-to-charge ratio which is analyzed can be continuously varied.

The output of the mass spectrometer shows a plot of relative intensity *vs* the mass-to-charge ratio (m/e). The most intense peak in the spectrum is termed the **base peak** and all others are reported relative to its intensity. The peaks themselves are typically very sharp, and are often simply represented as vertical lines.

(Image removed due to copyright reasons.)

Mass Spectrometry of Intact Proteins and Protein Fragments: The mass of intact proteins can be measured with an accuracy of $10,000 \pm 2$ Da. This ability provides a wealth of chemical information about protein chemical structure, including confirmation of amino acid sequence, detection and identification of chemical groups added post-translationally, and localization of proteolytic cleavage sites. The mass spectrum is like a *fingerprint* which reveals the structure of a compound either by skilled interpretation of the masses and abundances of *molecular ions* and their *fragment ions* or by comparison with records in a database (typically containing data for hundreds of thousands of prerecorded spectra).

This way, when you get the predicted sequence of small peptide fragments analyzed by mass spec you can later compare with full-length protein sequences in databases in order to identify your protein of interest. This is a very common method nowadays to determine proteins involved in the formation of complexes.

(A variant of mass spec is **Matrix-Assisted Laser Desorption Ionization (MALDI)**: MALDI is an ionization technique for large and/or labile molecules such as peptides, proteins, polymers, dendrimers, and fullerenes. This technique involves embedding analytes in a matrix which absorbs energy at the wavelength of the laser. The mechanism of the MALDI process, a combination of desorption and ionization, produces both positive and negative ions, usually in the forms of $(M+H)^+$, $(M+Na)^+$ and $(M-H)^-$. The technique also produces multiply charged ions, usually up to +3, as well as dimers, trimers, etc.)

Recombinant baculovirus expression in insect cells: The Baculovirus Expression Vector System (BEVS) is a convenient and versatile eukaryotic system for heterologous gene expression. Traditional methods involve heterologous protein expression in the bacteria *E. coli*. However, baculovirus expression provides correct folding of recombinant protein as well as disulfide bond formation, oligomerization and other important post-translational modifications (such as glycosylation). Consequently the overexpressed protein exhibits the proper biological activity and function (which sometimes cannot be attained by expressing the proteins in bacterial systems). The Baculovirus Expression Vector System is based on the introduction of a foreign gene into a nonessential region of the viral genome via homologous recombination with a transfer vector containing the cloned gene; an event that occurs in the co-transfected insect cells. The production of foreign protein is then achieved by infection of additional insect cell cultures with the resultant recombinant virus. The Baculovirus Expression Vector System usually employs a modified *Autographa californica* nuclear polyhedrosis virus (AcNPV) genome, and an appropriate transfer vector. The

diversity of AcNPV-based transfer vectors, combined with available *S. frugiperda* Sf9 and Sf21 insect cell lines, establish baculovirus expression as a preferred system for functional eukaryotic gene expression and the large-scale production of recombinant proteins.

(Images removed due to copyright reasons.)