de novo sequence assignment

de novo sequencing methods are traditionally employed in proteomic approaches in studies where the protein database is absent for an organism under study. It is in contrast to another popular peptide identification approach – "database search", which searches in a given database to find the target peptide. A clear advantage of de novo sequencing is that it works for both database and novel peptides.

Depending on the fragmentation methods used, different fragment ion types can be producedby MS/MS. The most widely used fragmentation methods are Collision-Induced Dissociation (CID) and Electron-Transfer Dissociation (ETD). CID produces mostly b and y-ions; and ETD produces mostly c and z-ions.

b and y-ions are the most common ion types, especially in the low-energy collision-induced dissociation (CID) mass spectrometers, since the peptide amide bone (CO-NH) is the most vulnerable and the loss of CO from b-ions.

Mass of b-ions = $\sum$ (residue masses) + 1 (H+)
Mass of y-ions = $\sum$ (residue masses) + 19 (H_2O+H+)

Amino acid sequence of the peptide can be interpreted from the tandem mass spectra of the peptides based on the mass difference between successive N-terminal (b-ions)/ and C-terminal (y-ions) ions.

Few factors to remember during sequence assignment:

Incorrect assignment of y and b ions is common.
Some fragment ions may be missing (such as b1 and y8).
Existence of other fragment ion types (such as the b3-NH3ion).
Existence of noise peaks in the spectrum.
The same or similar mass of some residues may cause ambiguity (I=L and K=Q).
The PTM (post-translational modifications) on the residues may contribute to the mass ambiguity, as well as complicate the peptide fragmentation pattern.

Amino acid influence

Basic residue at N-term strong b-ion series. (Arg>His>Lys)
Serine/threonine - x-18 ions frequently observed
Proline - N-terminal of residue(strong y-ion)
Aspartate- cleavage at c-terminal side (strong b-ion)
Val/leu/ile- C-terminal fragments(b-ion)
Gly/Ser- N-terminal fragmentation(y-ion)
N-G bond is very labile
His- side chain attack it's own c-term to give strong b-ion
Doubly protonated peptide fragment more easily than singly protonated forms of same peptide.

Large number of de novo sequencing software are available both independently and as a part of hybrid search platforms. Some examples below:

Peaks – a proprietary software, has de novo sequence search as an integral part of the search strategy. http://www.bioinfor.com/peaks/features/overview.html
Novor – denovo peptide sequencing
Direct tag – msms sequence tagging
Pepnovo+
Pnovo+
Lutefisk