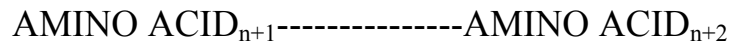


PROTEIN SEQUENCING

The overall strategy is to fragment a pure protein using several reagents with known specificity. That is, the reagent cleaves the protein at specific sites. Each fragment is then sequenced (by expensive machines). Lastly, the sequence of the original protein is assembled by looking for overlapping sequences (that's the fun part).



Amino Acid_{n+1} and amino acid_{n+2} are any two internal adjacent amino acids. The dashed line represents the peptide bond where hydrolysis will occur with the listed reagents.

Cyanogen bromide: Amino acid₁ = Meth

Trypsin: Amino acid₁ = Lys or Arg

Chymotrypsin: Amino acid₁ = Phe or Trp or Tyr

V8: Amino acid₁ = Asp or Glu

Thermolysin Amino acid₂ = Leu or Val or Ile

So, for example, if the original protein contains the amino acids Asp or Glu and is subjected to a V8 digest the resulting fragment(s) will have either Asp or Glu as their C-terminal amino acid, except for the fragment containing the original C-terminal amino acid of the starting peptide. If the original protein contains Leu, Val, or Ile and is subjected to a thermolysin digest the resulting fragment(s) will have either Leu, Val, or Ile as their N-terminal amino acid, again except for the fragment containing the original N-terminal amino acid of the starting peptide.

Note: carboxypeptidases are exopeptidases (in contrast to the above which are endopeptidases) in that they hydrolyze amino acids from the C-terminal end of the polypeptide.

Carboxypeptidase A does not cleave C-terminal Arg or Lys or residues that are next to Pro.

Carboxypeptidase B hydrolyzes only C-terminal Arg and Lys, but only if they are not preceded by Pro.