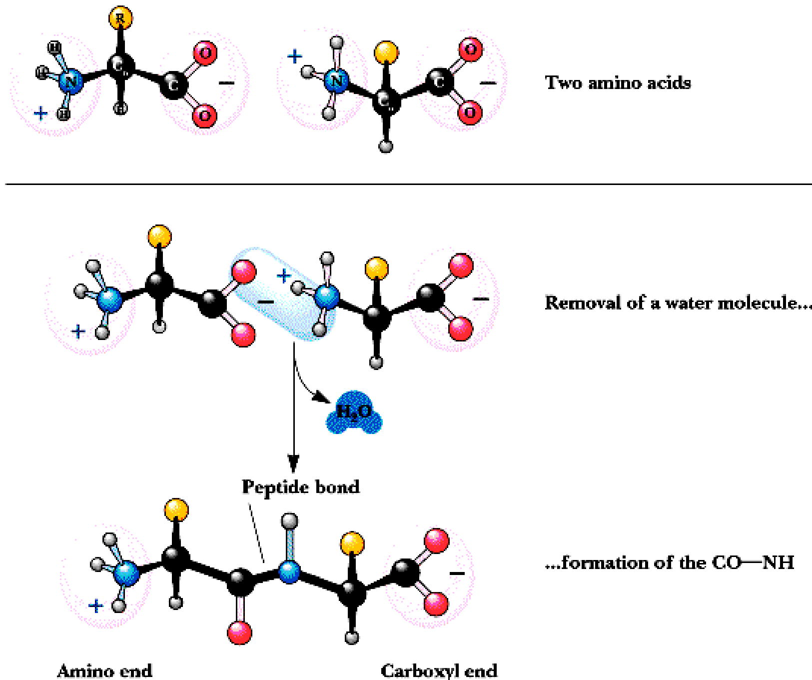
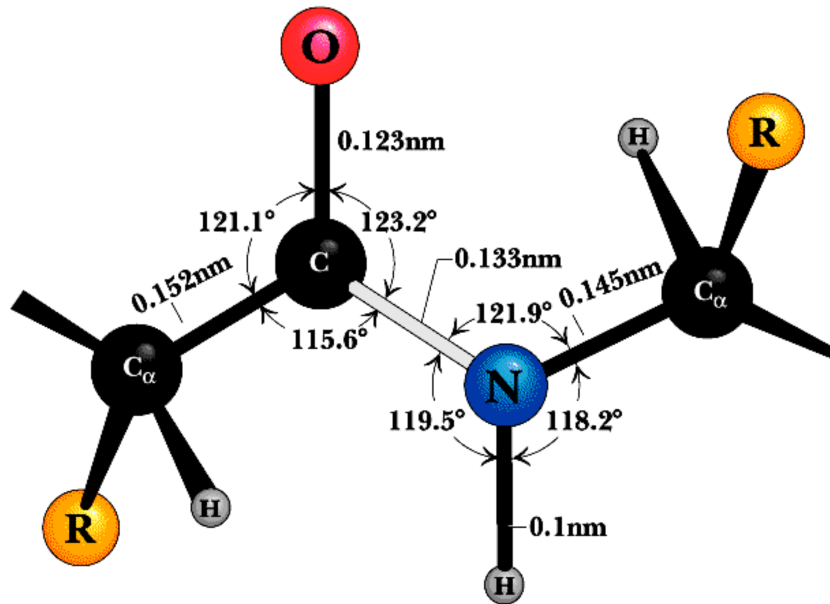
The Peptide Bond



- What allows amino acids to polymerize to form peptides and proteins is the unique covalent linkage called a peptide bond.
- The bond is the result of a head to tail condensation of the amino group of one amino acid and the carboxyl group of another. Formation of this bond results in the release of 1 mol of water per mol of peptide bond formed.
- A dipeptide contains two amino acids, a tripeptide contains three amino acids etc. In general, these structures are called oligopeptides. After 20 or so amino acids, oligopeptides begin to be called polypeptides
- Proteins are long polypeptides. Although the transition is vague, usually structures having molecular weights over 10,000 are called proteins.
- Reading Chap. 5 pp. 126-128, 137-150. Chap. 6 pp. 159-163.

Structure of the Peptide Bond



X-ray diffraction studies of crystals of small peptides by Linus Pauling and R. B. Corey indicated that the peptide bond is rigid, and planer.

Pauling pointed out that this is largely a consequence of the resonance interaction of the amide, or the ability of the amide nitrogen to delocalize its lone pair of electrons onto the carbonyl oxygen.

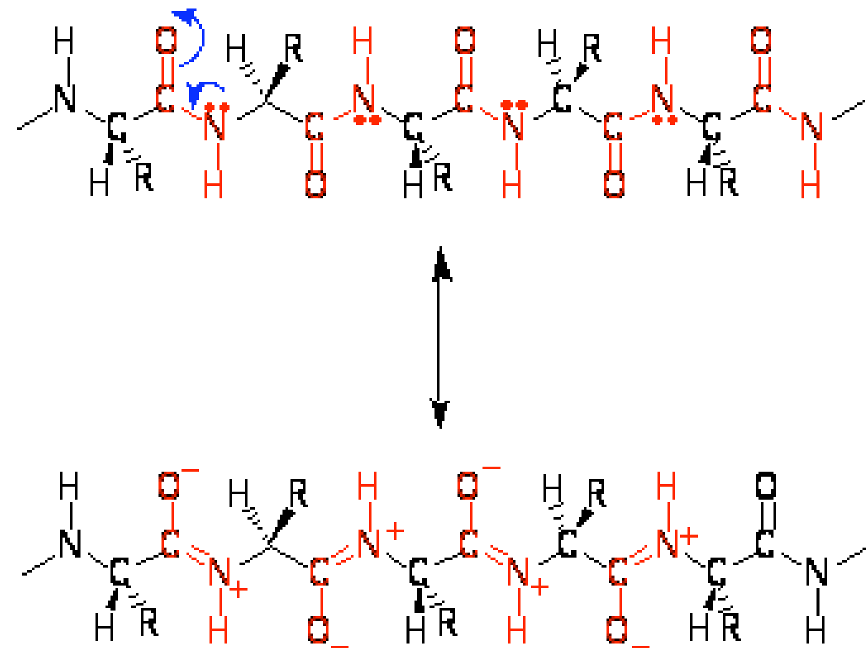
Because of this resonance, the C=O bond is actually longer than normal carbonyl bonds , and the N–C bond of the peptide bond is shorter than the N–Cα bond.

Notice that the carbonyl oxygen and amide hydrogen are in a trans configuration, as opposed to a cis configuration. This configuration is energetically more favorable because of possible steric interactions in the other.

The Polarity of the Peptide Bond

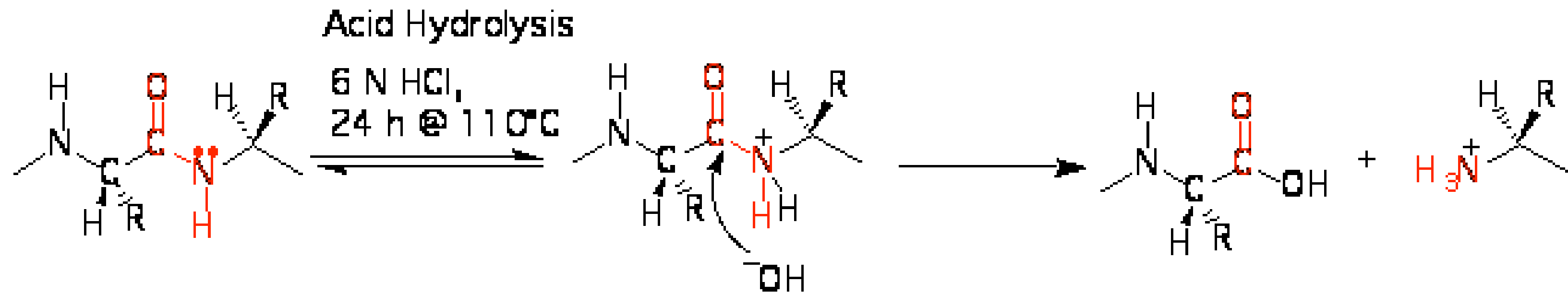
The peptide bond is usually portrayed as a single bond between the carbonyl carbon and the amide nitrogen. Normally, this should allow free rotation about that bond. However, notice that the nitrogen has a lone pair of electrons, which are adjacent to a carbon-oxygen bond. Therefore, a reasonable resonance structure can be drawn with a **double** bond linking the carbon and nitrogen, and which results in a negative charge on the oxygen and a positive charge on the nitrogen. Remember, when writing resonance structures that atoms cannot move, and the total charge must remain the same.

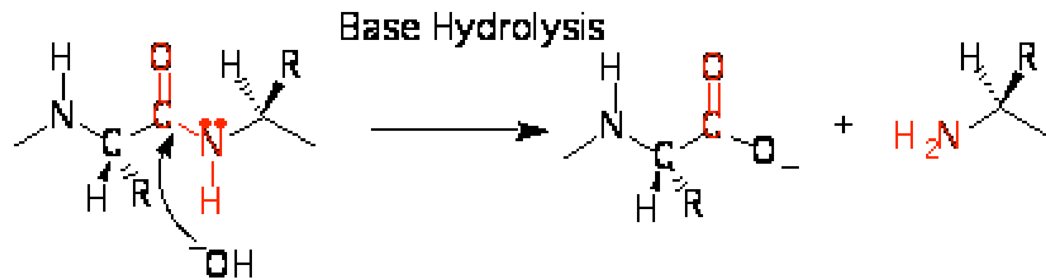
The resonance structure prevents rotation around the peptide bond. The real structure of course is a weighted hybrid of these two structures. Therefore, the question is how significant is the resonance structure in depicting the true electron distribution. It is known that the peptide bond has approximately 40% double-bond character. It is therefore rigid.

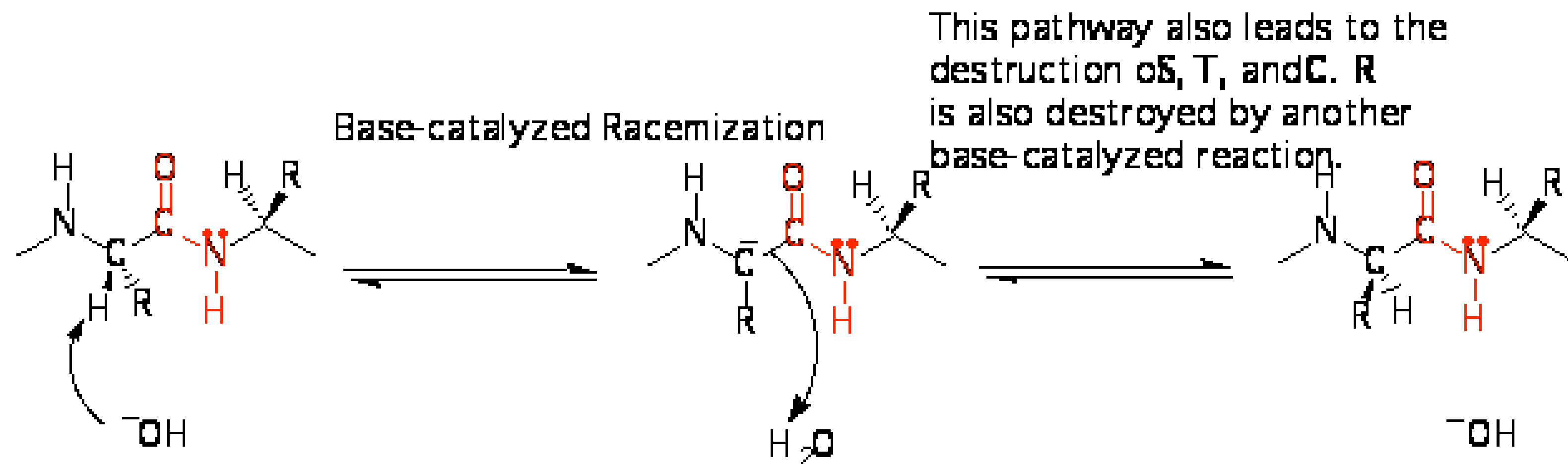


Charges give the peptide bond a permanent dipole. Because of the resonance, the oxyge has a -0.28 charge, while the nitrogen bears +0.28 charge.

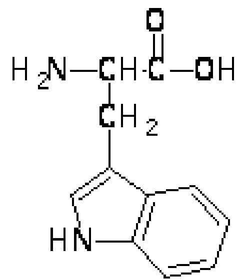
Hydrolysis of the Peptide Bond





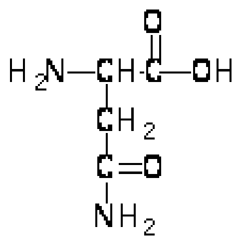


Acid Hydrolysis of Peptide Bond

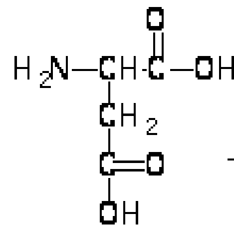


W

W is destroyed during acid hydroly:

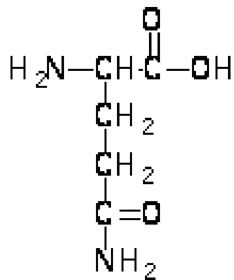


N

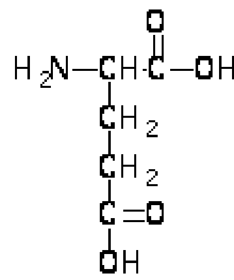


D

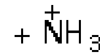
Acid hydrolysis destroys several different amino acids. Notably, tryptophan is totally destroyed. Cysteine is destroyed, and asparagine and glutamine are hydrolyzed to the corresponding carboxylic acids with subsequent release of ammonia.



Q

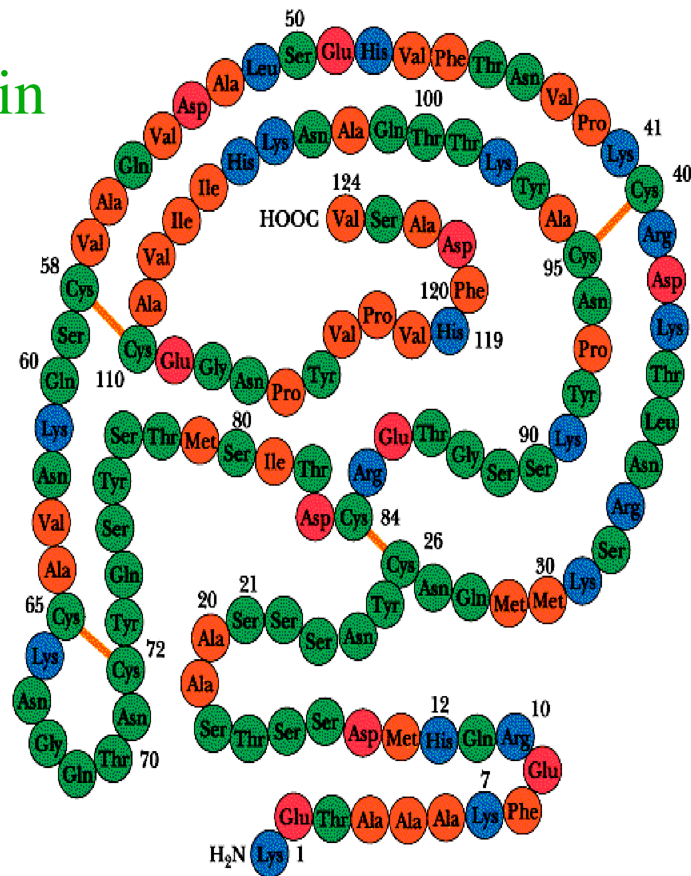


E



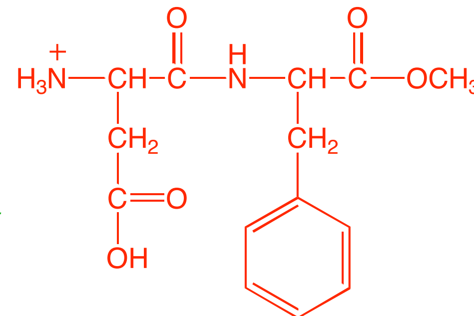
Protein Sequences

Insulin



- The primary structure of a protein represents the linear arrangement of amino acids via peptide bonds. By convention, the sequence is read from N-terminus to C-terminus
- Every molecule of a particular protein will have the exact same amino acid sequence. Different proteins have different lengths and amino acid variations.
- In general, the primary sequence of a protein will dictate its higher order structures–such as secondary, tertiary, and quaternary structures–as well as its function.
- Biologically relevant peptides range in size from two amino acids to proteins containing thousands of amino acids. NutraSweet, for example, is the dipeptide L-aspartyl-L-phenylalanine methyl ester.

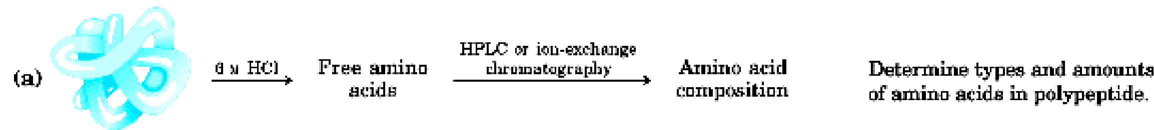
NutraSweet

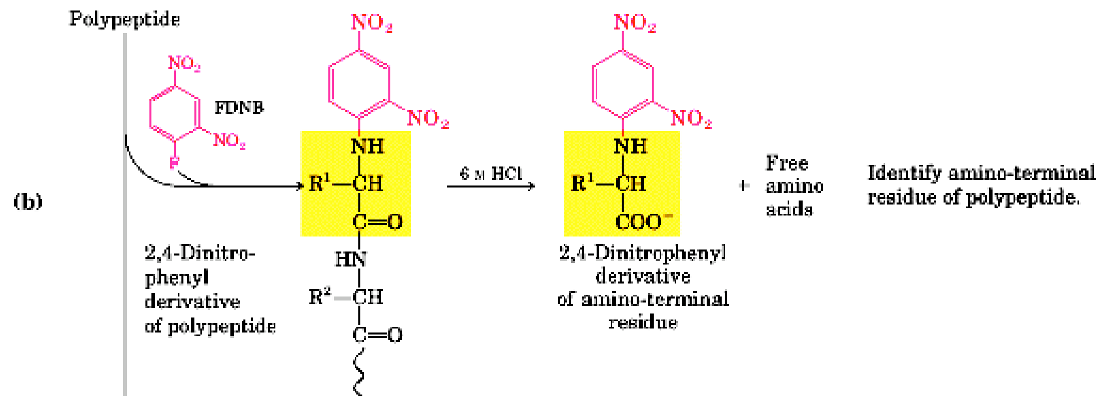


Protein Sequencing Strategies (Chemical)

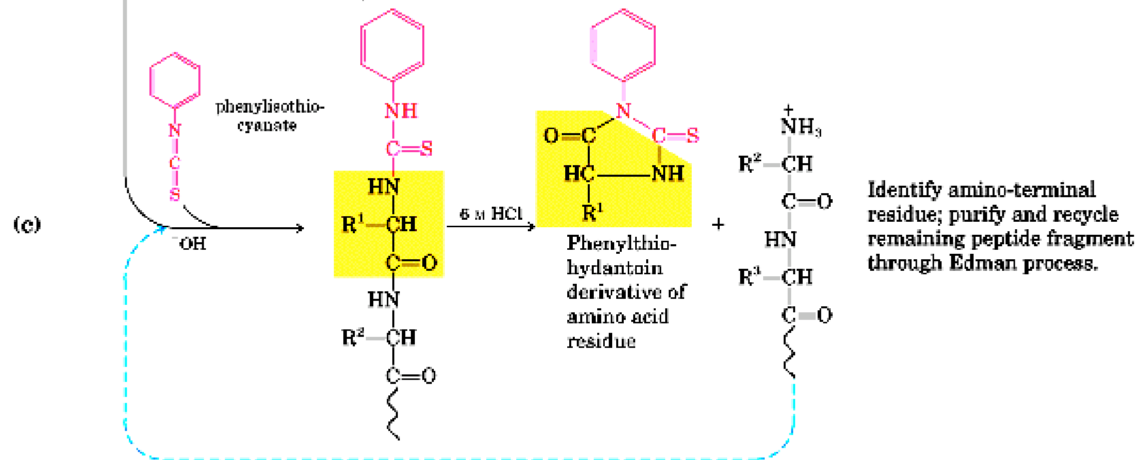
- Before the advent of modern DNA technology, the sequencing of proteins was very laborious and frequently inaccurate. In fact, many thought that it would be an insurmountable task.
- In 1953, Frederick Sanger worked out the sequence of the amino acid residues that comprise the polypeptides of the hormone insulin, for which he received the Nobel Prize. Note that Frederick Sanger received another Nobel Prize later for developing a method for sequencing DNA.
- Sequencing by chemical methods
  - Separation of the polypeptide chains of multimeric proteins.
  - All disulfide bonds must be cleaved.
  - Determine the amino acid composition of each of the chains.
  - Determine the N-terminal and C-terminal residues of each chain
  - Cleave each chain into smaller more manageable fragments, and determine the amino acid composition and sequence. Repeat using methods that cleave the chain in different positions.
  - Reconstruct the entire amino acid sequence by piecing together the information gained from the fragments.
  - Locate the positions of disulfide bonds.

Strategy for Protein Sequencing





1-fluoro-2,4-dinitrobenzene
FDNB (Sanger's Reagent)

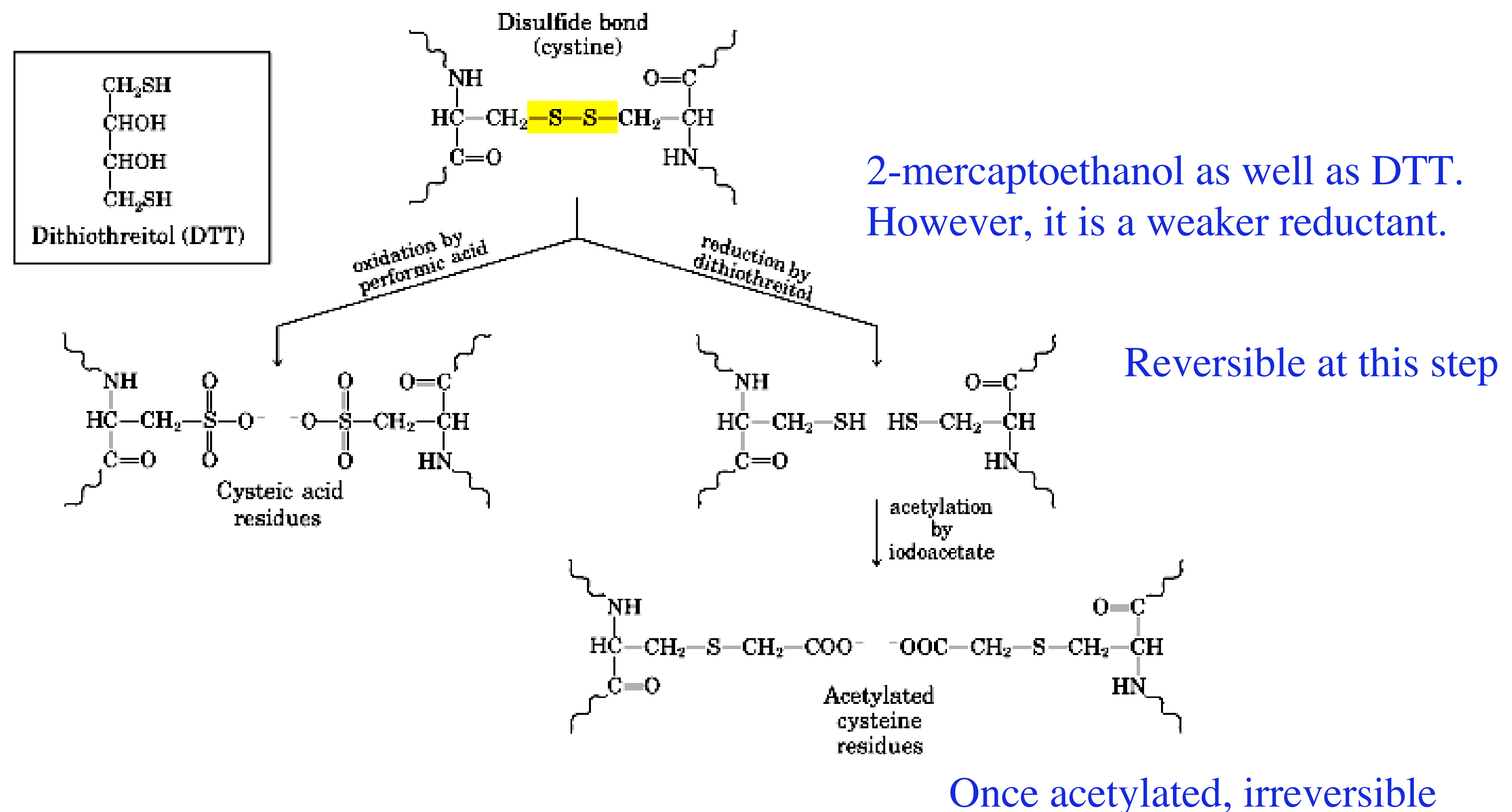


Edman Degradation

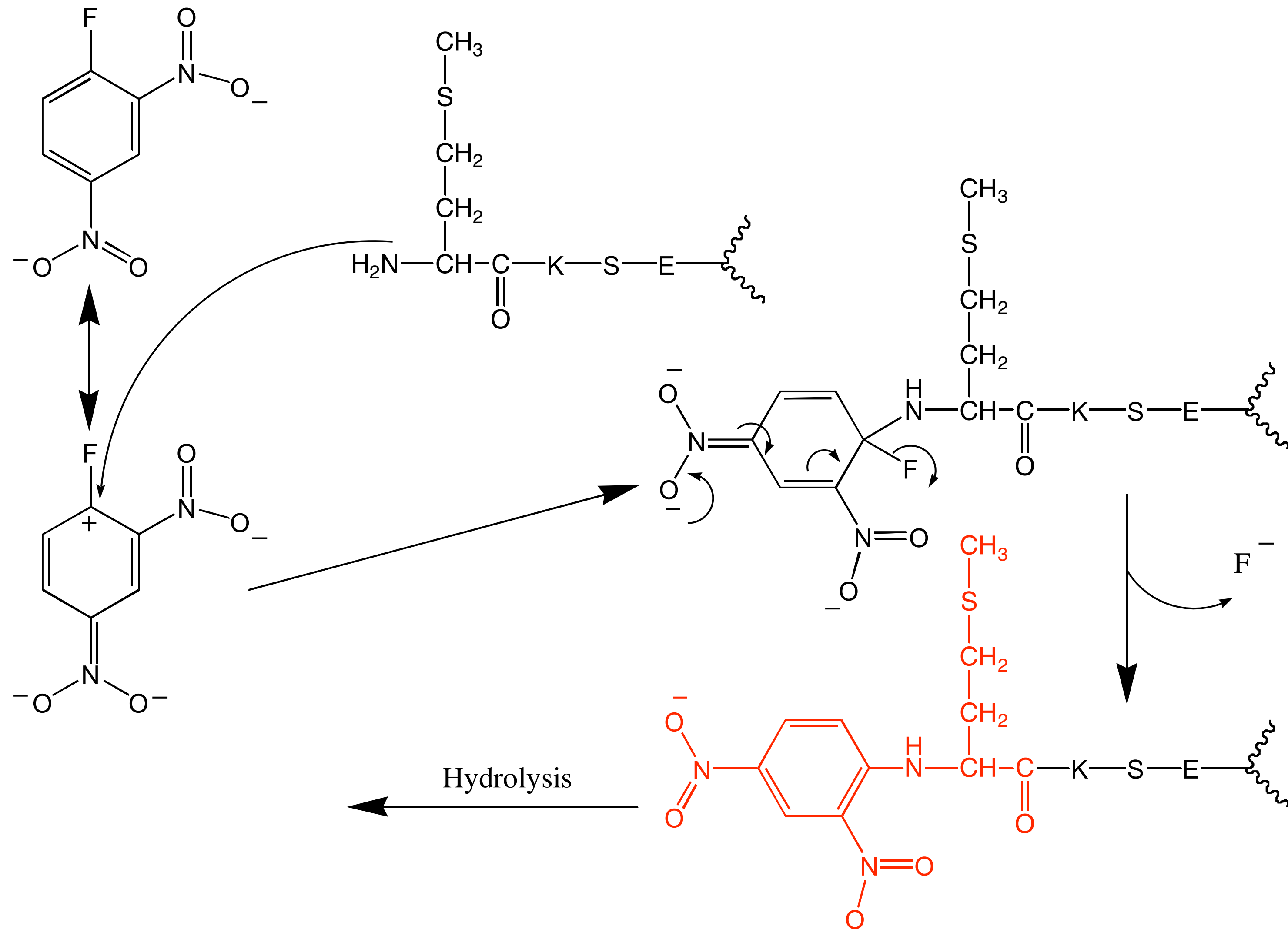
Disulfide-Bond Cleavage

Polypeptide chains of multimeric proteins are usually held together by noncovalent forces, and can be dissociated by exposure to pH extremes, chaotropic agents (urea, guanidinium hydrochloride), or high salt concentrations.

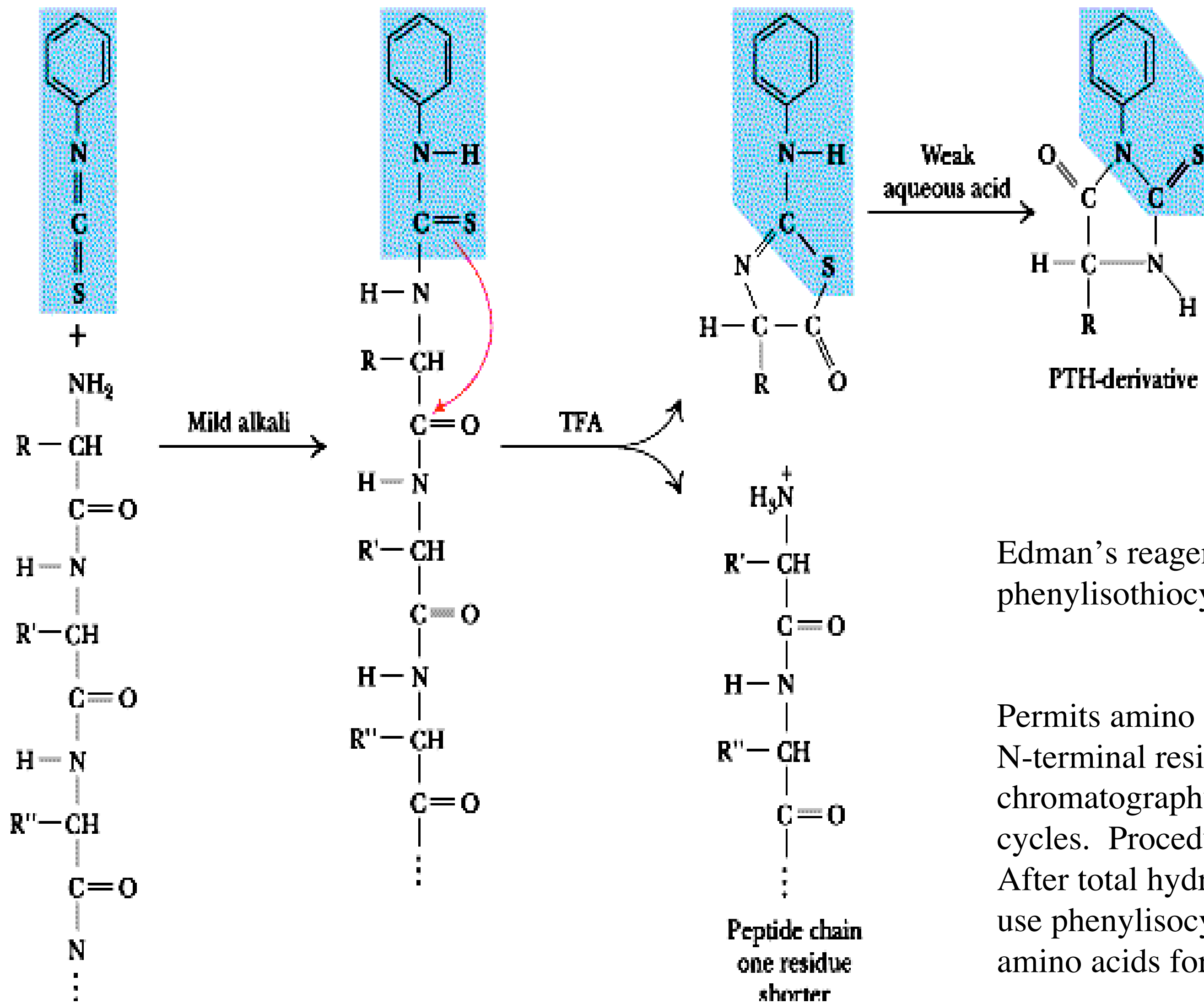
Disulfide bond cleavage should be carried out in a way that prevents them from reforming. A number of methods exist.



1-Fluoro-2,4-nitrobenzene



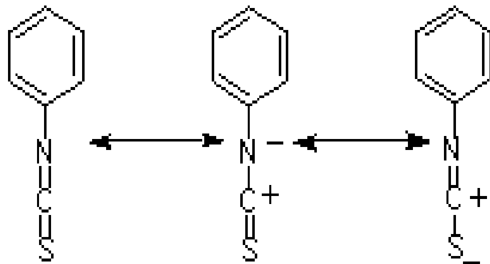
N-terminal Analysis



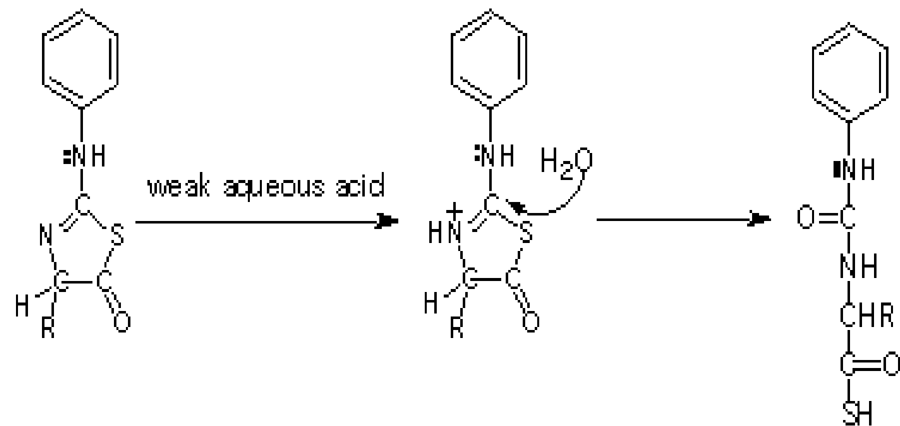
Edman's reagent - phenylisothiocyanate.

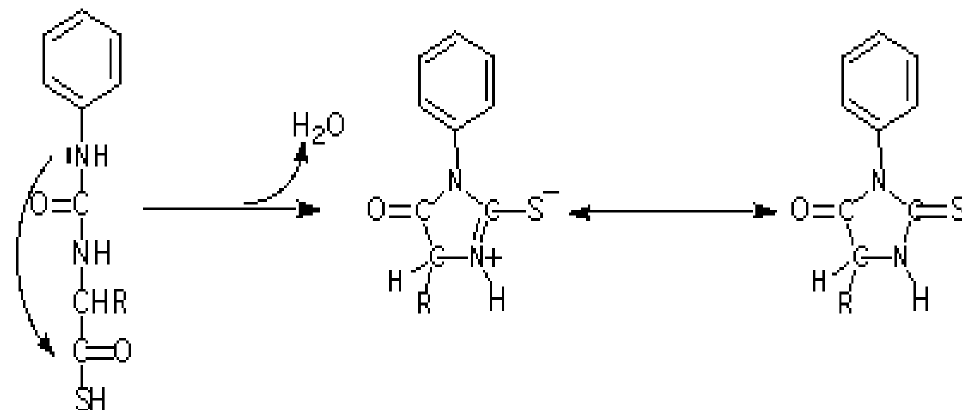
Permits amino acid determination of N-terminal residue after chromatography. Can do repeated cycles. Procedure is now automated. After total hydrolysis of protein, can use phenylisocyanate to derivatize amino acids for analysis.

Chemistry of Edman Degradation

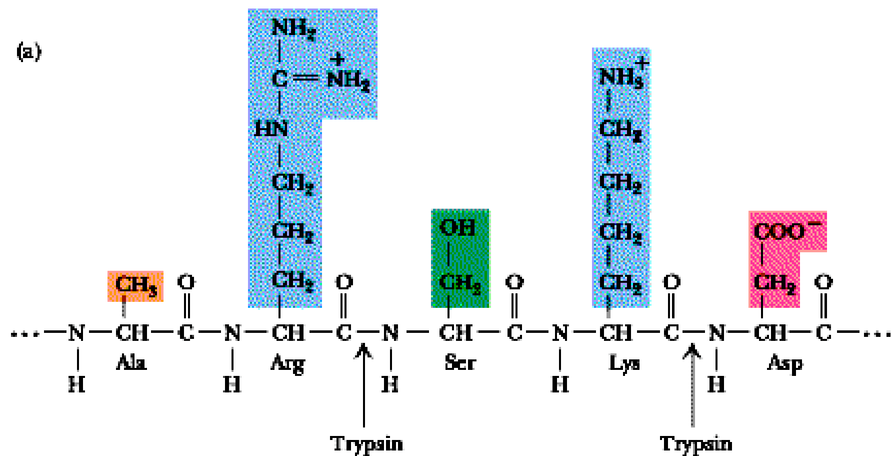


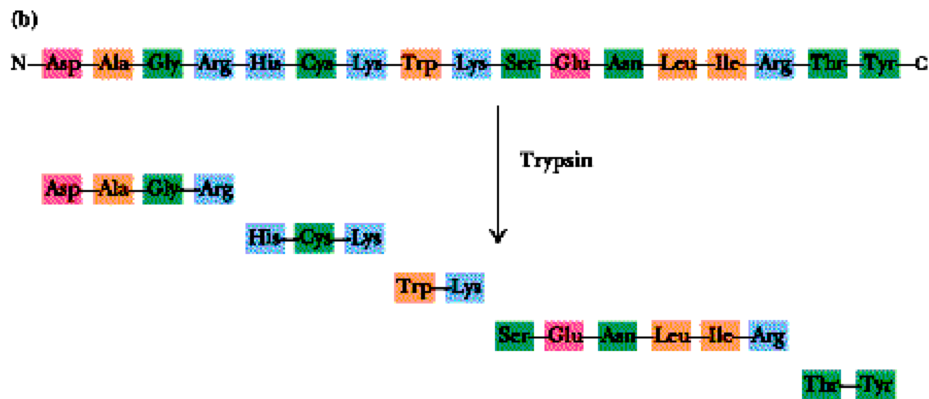
Some Resonance forms of Edman's Reagent. Note that the carbon has positive character, making it electrophilic





C-terminal Analysis and Peptide Fragmentation





The C-terminal residue of proteins is usually determined by an enzymatic reaction. Carboxypeptidases are enzymes that cleave amino acids from the C-terminal end of polypeptides in a successive fashion. These are called exopeptidases. Four are generally recognized:

Carboxypeptidase A cleaves the C-terminal peptide bond of all residues except P, R, and K.

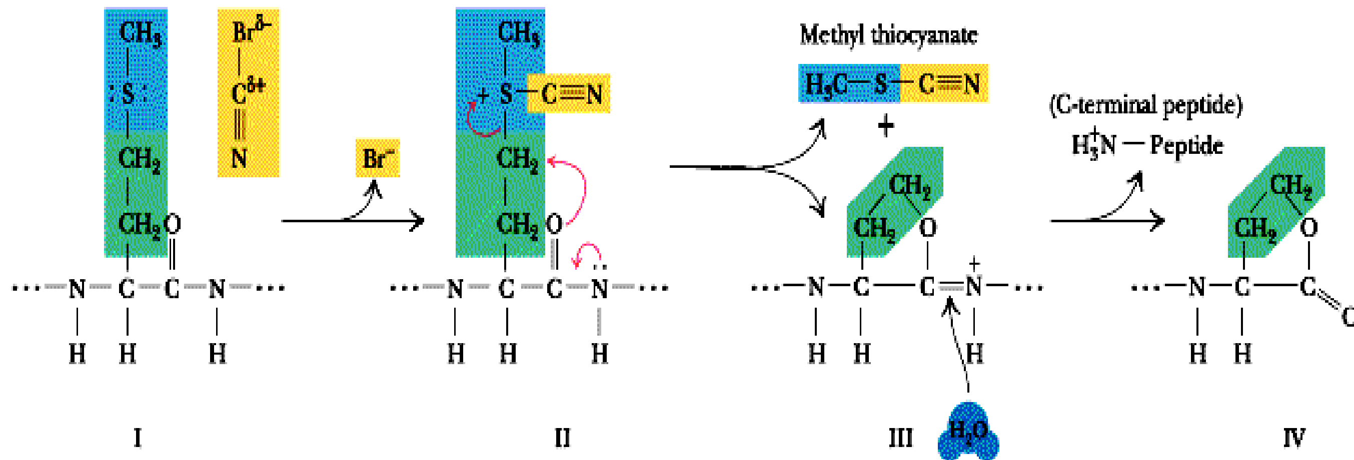
Carboxypeptidase B cleaves only when R and K are at the C terminus

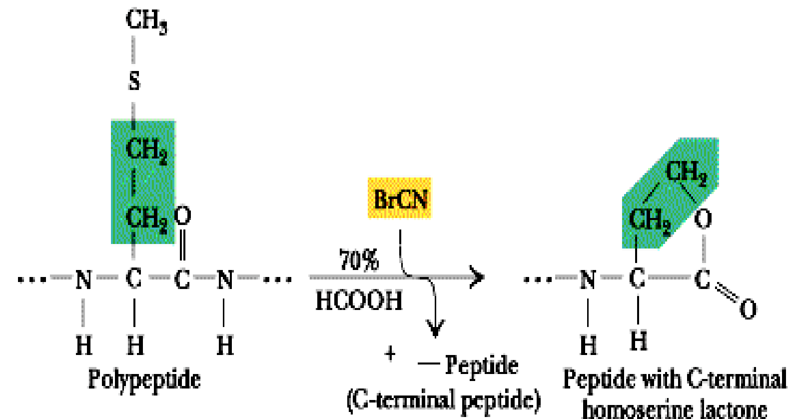
Carboxypeptidases C and Y act on any C-terminal residue.

Endopeptidases cleave within a polypeptide chain. *Chymotrypsin* cleaves **after** F, Y, and W, and also L, but less well. *Trypsin* cleaves after R or K residues. *Clostripain* cleaves only after R, while Lys-C cleaves only after K. Also, *staphylococcal protease* cleaves after D or E. *Papain* is relatively nonspecific, and is used in laundry detergent, contact lens cleaner, and meat tenderizer.

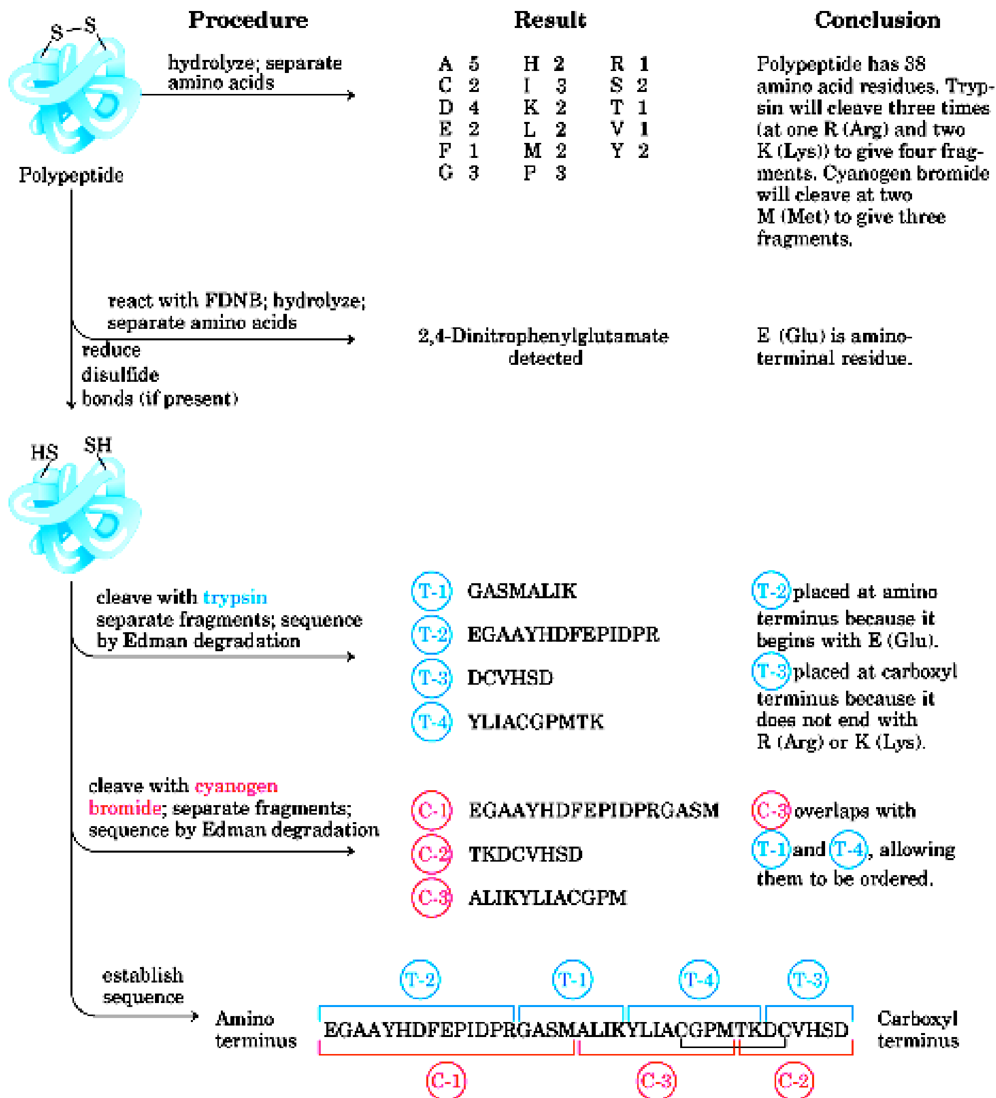
Chemical Cleavage of the Peptide Backbone

Cyanogen bromide is a highly selective cleavage method it cleaves only after methionine residues. Produces a modified C-terminus at all cleavage sites containing a homoserine lactone residue. Only the original C-terminus will not contain this modification.





Strategy in Action



Here is a hypothetical polypeptide. Notice the sequence of events that enable the elucidation of its primary sequence.

To locate position of disulfides, cleave the protein with a protease in the presence and absence of prior treatment with reducing agents. After separation of the resulting peptides, it will be noticed that in the presence of the reducing agent, two new peptides will appear, and one peptide will be missing. Sequence the two new peptides.

How About This Problem?

- A cyanogen bromide fragment has been purified from a digest of a certain protein. Consider the following information. The compositions shown in parentheses are those obtained following complete acid hydrolysis in 6 M HCl, 110°C, for 24 h.
- (A) Complete acid hydrolysis gives R, E, 2 G, homo-S, L, K, F, S, V
- (B) Edman degradation produced the phenylthiohydantoin of V
- (C) After trypsin digestion, the amino acid composition of peptides was the following
  - (1) R, E, G, V
  - (2) G, K, F, S
  - (3) homo-S, L
- (E) Edman degradation
  - V-E for peptide 1 and F-S for peptide 2
- (F) Reaction with 2,3-butanedione, followed by tryptic digest
  - R, E, 2 G, K, F, S, V
  - Homo-S, L

  - **What is the peptide sequence? What amino acid does 2,3-butanedione interact with?**

Amino Acid Analysis

- Amino acid analysis will only give the ratio of the various amino acids. If the molecular weight of the protein is known as well as the exact amount of protein hydrolyzed, the number of each residue can be determined. This procedure does not give the primary sequence, only the ratio of the various amino acids that make up the primary sequence

- For instance, if you hydrolyze 2 mg of a protein having a molecular weight of 50,000 g/mol, and get 2 μmol of alanine from amino acid analysis, then you can calculate the number of alanine residues. 0.002 g x (1 mol protein / 50,000 g) = 40 nmol
  - 2 μmol / 40 nmol = 50

  - Similarly, you can determine the extinction coefficient for a particular protein by amino acid analysis if you know how many of a particular residue (for example, alanines) the protein has. If 1 mL of the above protein (2 mg) gave an absorbance of 2, the molar absorptivity would be 50,000 M^{-1}•cm^{-1}.