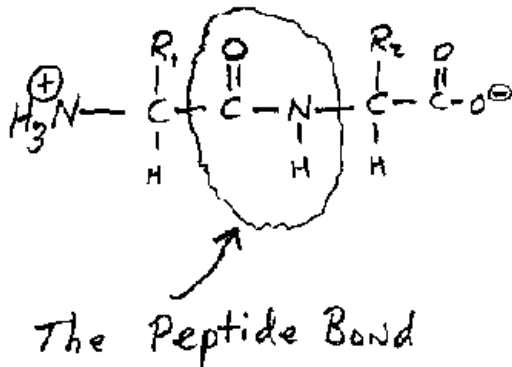


BL/CH 401 Lecture #5

Protein Covalent Structure (Protein Primary Structure)

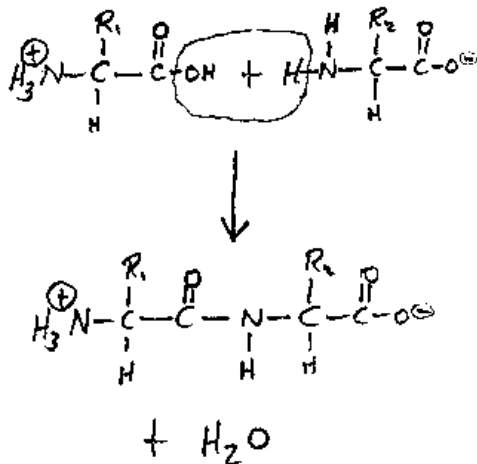
I. Peptide Bonds, Peptides and Proteins

Proteins are sometimes called Polypeptides, since they contain many **Peptide Bonds**



{*Figure 1*}

The **peptide bond** is an **amide bond**



{*Figure 2*}

Water is lost in forming an amide bond.

Structural Character of Amide Groups: Understanding the chemical character of the amide is very important, since the peptide bond of proteins is an amide bond.

Amino Acid = Gly; dipeptide = Gly-Ala; tripeptide = Gly-Ala-Ser

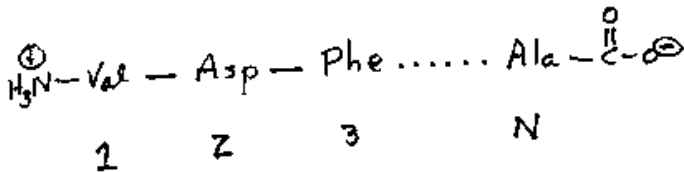
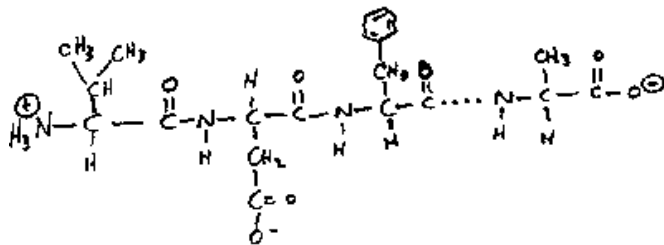
Peptides = Mini-Proteins



A pentapeptide -- GlyAlaSerPheGln

1st amino acid is always written on the left and called the Amino terminal, since it is always the only amino acid of the peptide with a free alpha-amino group. Last amino acid is always written on the right and called the Carboxyl terminus, since it is always the only amino acid of the peptide with a free alpha-carboxylic acid group.

Full structure of a Protein



{*Figure 5*}

Amino acid sequence of a Protein.

II. Amino Acid Composition

Amino Acid analysis yields a protein's Amino Acid Composition, in other words the amounts of each amino acid in the protein. All proteins have unique amino acid compositions.

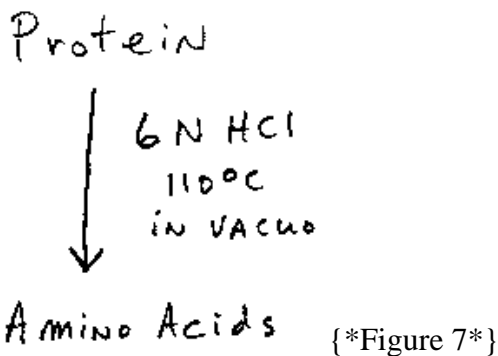
Amino acid compositions of a few proteins (see list proteins below table)
Values for Amino Acids are in Number of Residues per Molecule of Protein

Amino Acid \ Protein	A	B	C	D	E	F	G	H	I
Nonpolar									
Ala (A)	12	6	9	1	3	27	15	12	27
Val (V)	13	3	7	1	4	13	18	6	14
Leu (L)	15	6	8	2	6	18	18	8	38
Ile (I)	9	8	4	0	1	0	0	6	12
Pro (P)	11	4	4	0	1	7	7	2	2
Met (M)	1	3	0	1	0	2	1	2	2
Phe (F)	10	3	2	2	3	7	8	3	4
Trp (W)	2	1	1	1	0	1	2	6	0
Polar, Uncharged									
Gly (G)	13	13	6	1	4	7	13	12	14
Ser (S)	28	2	7	4	3	11	5	10	27
Thr (T)	19	7	8	3	1	9	7	7	15
Tyr (Y)	9	5	4	2	4	3	3	3	8
Cys (C)	5	2	5	0	6	1	2	8	15
Asn (N)	10	5	2	1	0	4	6	13	17
Gln (Q)	16	2	4	3	0	1	3	3	25
Polar, Negative Charge									
Asp (D)	10	3	11	3	3	8	7	8	17
Glu (E)	10	8	9	0	7	4	8	2	26
Polar, Positive Charge									
Lys (K)	13	18	4	1	1	11	11	6	18
Arg (R)	6	2	1	2	9	3	3	11	11
His (H)	2	3	1	1	2	10	9	1	2
Total Residues	214	104	97	29	54	141	146	129	312

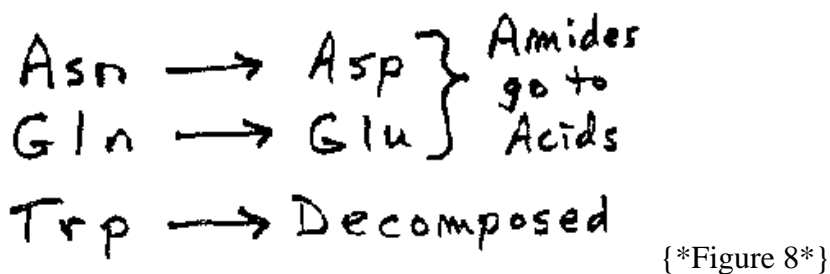
List of Proteins Shown in Amino Acid Composition Table:

- A. Antibody - Human Bence-Jones Kappa (antibody light chain)
- B. Human Cytochrome c (electron transport protein)
- C. Spinach Ferredoxin (electron transport protein)
- D. Pig Glucagon (protein hormone)
- E. Bovine Insulin (protein hormone)
- F. Human/Gorilla Hemoglobin alpha chain (oxygen transport protein)
- G. Human/Gorilla Hemoglobin beta chain (oxygen transport protein)
- H. Chicken Lysozyme (enzyme)
- I. Sheep Wool (structural protein)

Free amino acids are obtained from proteins by strong acid hydrolysis:



3 of the standard 20 AAs are lost during Acid Hydrolysis treatment:



The amide AAs, Asn & Gln, are converted to their acids, Asp & Glu. Trp is simply unstable under acid hydrolysis conditions and is destroyed. Despite the loss of 3 AAs, 6 N HCl and 110 C are the conditions of choice for protein hydrolysis. The amounts of the remaining 17 AA are determined by Amino Acid Analysis with an **Amino Acid Analyzer** (an analytical instrument).

III. Amino Acid Sequencing of Proteins and Peptides

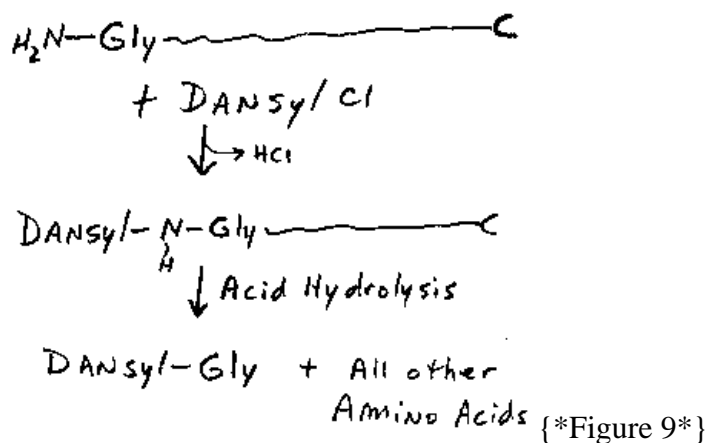
General Approach to Determining the Amino Acid Sequence of a Protein:

1. Purify protein to homogeneity (See Lecture 6 & 7)
2. Determine AA composition (See #2 Above)
3. Reduce -S-S- bonds to -SH groups and block
4. Determine N- and C-terminal amino acids

Determination of Amino and Carboxyl Terminal Amino Acids of Proteins and Peptides

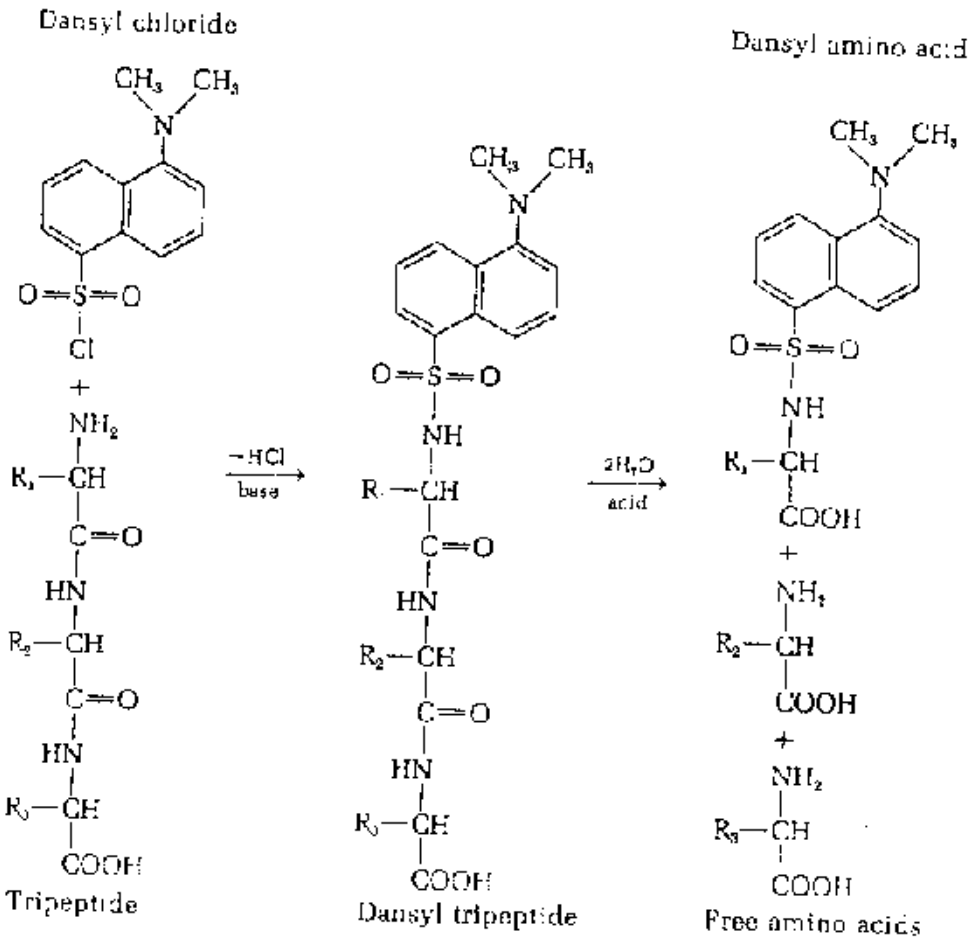
Determine N- and C- terminal amino acids, so that you can easily recognize the N- and C-terminal peptides after the protein is cleaved into fragments.

N-terminal Determination: best method is the **dansyl chloride**:



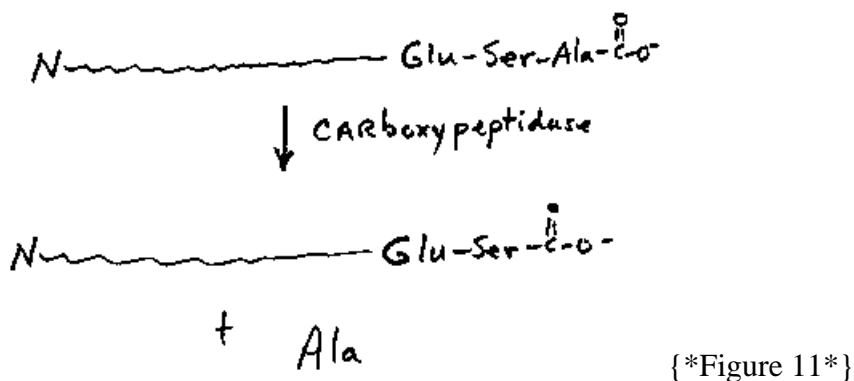
General scheme for N-terminal amino acid determination.

Dansyl Chloride yields a fluorescent derivative of N-terminal amino acid so method is very sensitive! Chemical structures for Dansyl chloride method:



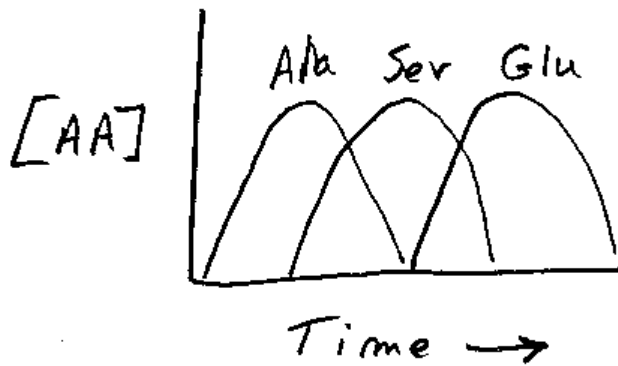
{*Figure 10*} Dansyl method for N-terminal AA Determination.

C-terminal Determination: best method is enzymatic degradation with **carboxypeptidase**:



General scheme for C-terminal amino acid determination.

Usually several AAs are released from the C-terminal by carboxypeptidase, so a time-course determination is done to see which AA is released first:



{*Figure 12*}

Carboxypeptidase treatment of a peptide: Xxx-----GluSerAla

5. Cleave protein into peptides

Cleavage of the Protein into Smaller Peptides

So after you have the N- and C-termini determined, you cleave the protein into peptide fragments using:

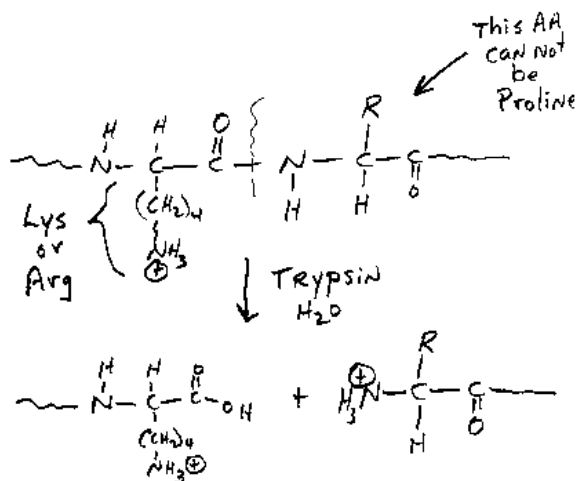
1. Hydrolytic Enzymes (isolated from animal digestive system):

A. Trypsin cleaves at Lys and Arg

B. Chymotrypsin cleaves at Aromatic AAs (Phe, Tyr & Trp)

2. Specific chemicals like Cyanogen Bromide (CNBr) which cleaves at Met

Trypsin catalyzes peptide bond hydrolysis with water:

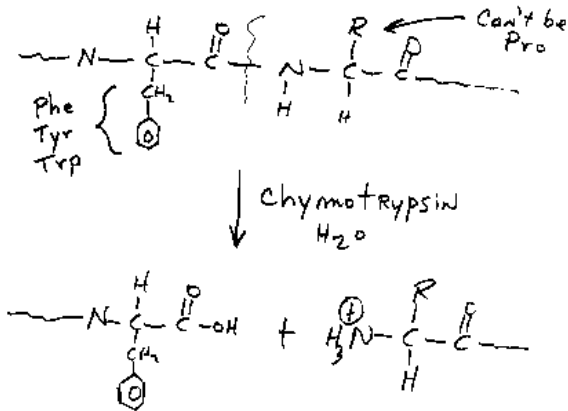


{*Figure 13*}

Important to note that the AA following the target Lys/Arg can not be Pro. Why does trypsin not catalyze breakage of Lys-Pro and Arg-Pro bond? Remember that Pro has an unusual alpha-

amino group and an unusually inflexible structure hence, Lys-Pro & Arg-Pro peptide bonds do not bind to trypsin's catalytic site well!

Chymotrypsin catalyzes peptide bond hydrolysis with water just like trypsin:



{*Figure 14*}

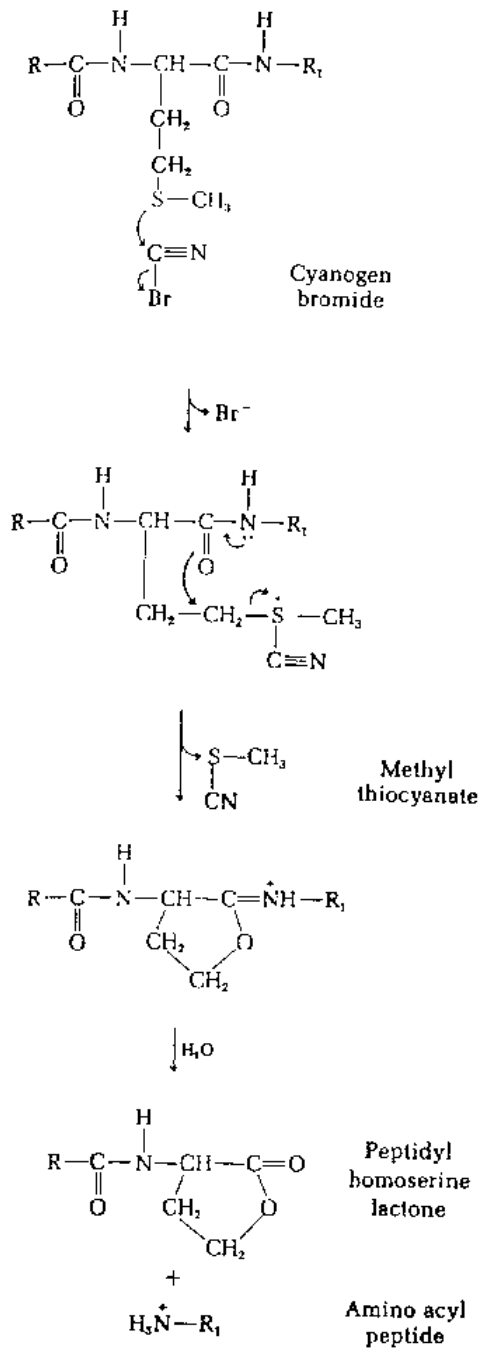
Important to note that the AA following the target Phe/Tyr/Trp can not be Pro. Why does chymotrypsin not catalyze breakage of Phe-Pro, Tyr-Pro and Trp-Pro bond? For same reason that trypsin does not (see discussion above)!

For trypsin and Chymotrypsin the number of peptides equals the number of target AAs plus one more.

For example: If protein has 4 Lys and 3 Arg then get 8 peptides. Four will have Lys at C-terminal, three will have Arg at C-terminal, and the carboxyl-terminal peptide will not contain Lys or Arg.

Cyanogen Bromide (CNBr) cleaves at Met and...converts the Met into another amino acid called Homoserine lactone (HSL).

If a protein contains 3 Met, then 4 peptides will be found - 3 with C-terminal HSL and one containing no HSL.

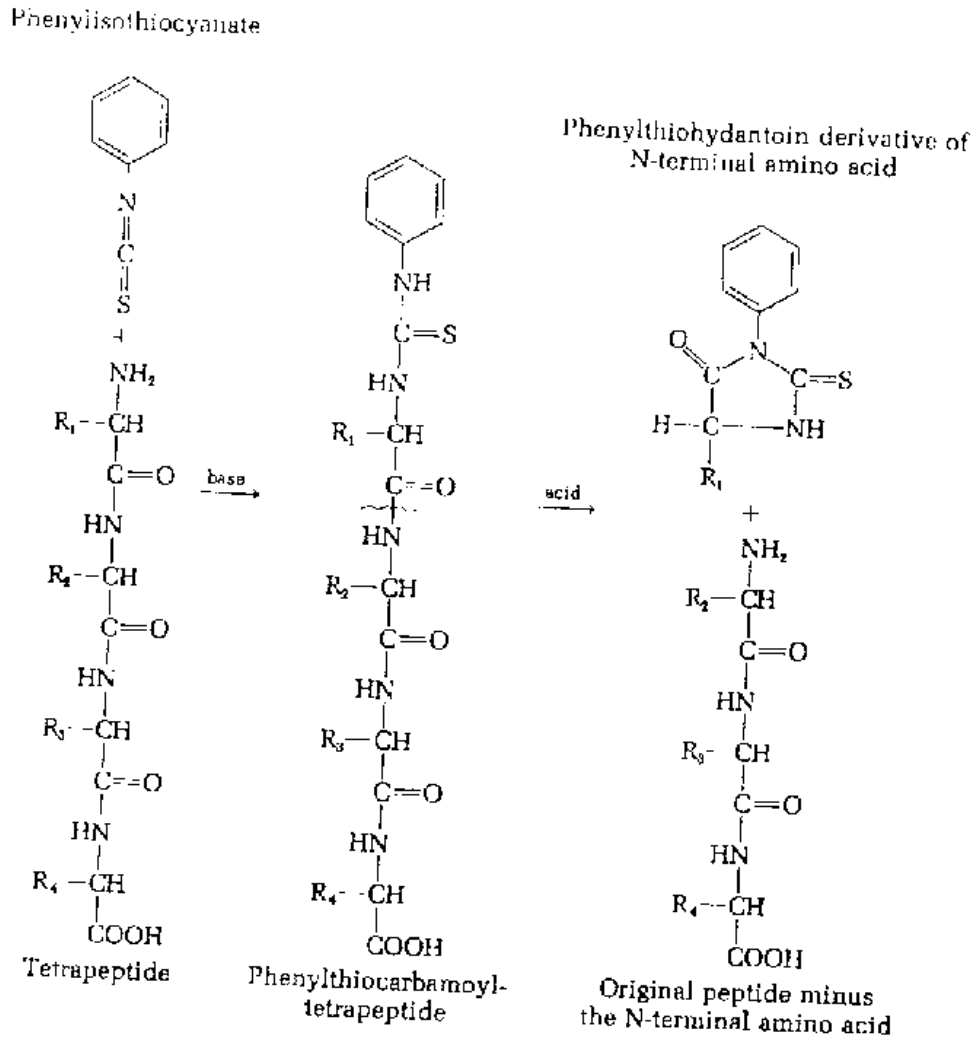


{*Figure 15*} CNBr cleavage of peptide

6. Separate peptides using methods for protein purification. (see lectures 6 & 7)
7. Determine amino acid sequence of the peptides

Edman Method for Peptide Sequencing

Finally, after each peptide is separated into a pure form from the mixture of peptides generated by the cleavage method, the amino acid sequence of each peptide is determined using the **Edman Reagent: Phenylisothiocyanate**. It works by cleaving N-terminal AA off and leaving rest of the peptide intact.



{*Figure 16*}

The **Edman method** can be repeated on the peptide remaining after the reaction, so this is a very effective method for obtaining AA sequence information.

In fact the whole sequence of a peptide containing many AAs can be determined by Edman method.

8. Use a second method to cleave the protein into peptides
9. Separate and sequence the second set of peptides

10. Deduce the entire amino acid sequence by identifying overlapping peptide sequences

After all peptides from one cleavage method have been sequenced, then the protein is cleaved by a second method and peptides separated and sequenced. So one sequences the protein twice. Overlapping sequences are identified so that the order of the peptides in the sequence can be deduced and consequently the entire amino acid sequence of the protein.

AMINO ACID SEQUENCE DETERMINATION HELP SHEET

1. Acid hydrolysis of a protein is used to determine its amino acid composition:

only 17 AAs survive acid hydrolysis

Asn & Gln are converted to Asp & Glu

Trp is destroyed

AA composition helps you find the number of cleavage sites for an enzyme or CNBr

For example, if a peptide contains 2 Lys & 1 Arg, then Trypsin digestion may yield 4 peptides

2. N-terminal residues can be determined with Dansyl-Cl method, which reacts with free amino groups. After acid hydrolysis, the Dansyl-AA can be determined.

3. C-terminal residues can be determined by enzymatic digestion with the enzyme carboxypeptidase, which releases amino acids sequentially.

4. Enzymatic and Chemical methods for peptide cleavage into smaller peptides are:

A. Enzymatic -- Trypsin -- cleaves on carboxyl side of Lys & Arg (unless Pro is next AA)

GlyArgSerLysMetAlaTrpTyrProPheGlyLeuThr

↓ Trypsin Digestion

GlyArg + SerLys + MetAlaTrpTyrProPheGlyLeuThr

B. Enzymatic -- Chymotrypsin -- cleaves on carboxyl side of Aromatic AAs (Phe, Tyr, Trp)
(unless Pro is next AA)

GlyArgSerLysMetAlaTrpTyrProPheGlyLeuThr

↓ Chymotrypsin Digestion

GlyArgSerLysMetAlaTrp + TyrProPhe + GlyLeuThr

C. Chemical -- CNBr cleaves on carboxyl side of Met & converts it to Homo-Ser-Lactone (HSL)

GlyArgSerLysMetAlaTrpTyrProPheGlyLeuThr

↓ CNBr Treatment then mild acid

GlyArgSerLysHSL + AlaTrpTyrProPheGlyLeuThr

5. Edman degradation removes AA from the N-terminus, 1 at a time and can be repeated on the residual peptide remaining after the first AA is removed:

GlyArgSerLysMetAlaTrpTyrProPheGlyLeuThr

↓ Treatment with PITC then mild acid (Edman)

PTH-Gly + ArgSerLysMetAlaTrpTyrProPheGlyLeuThr