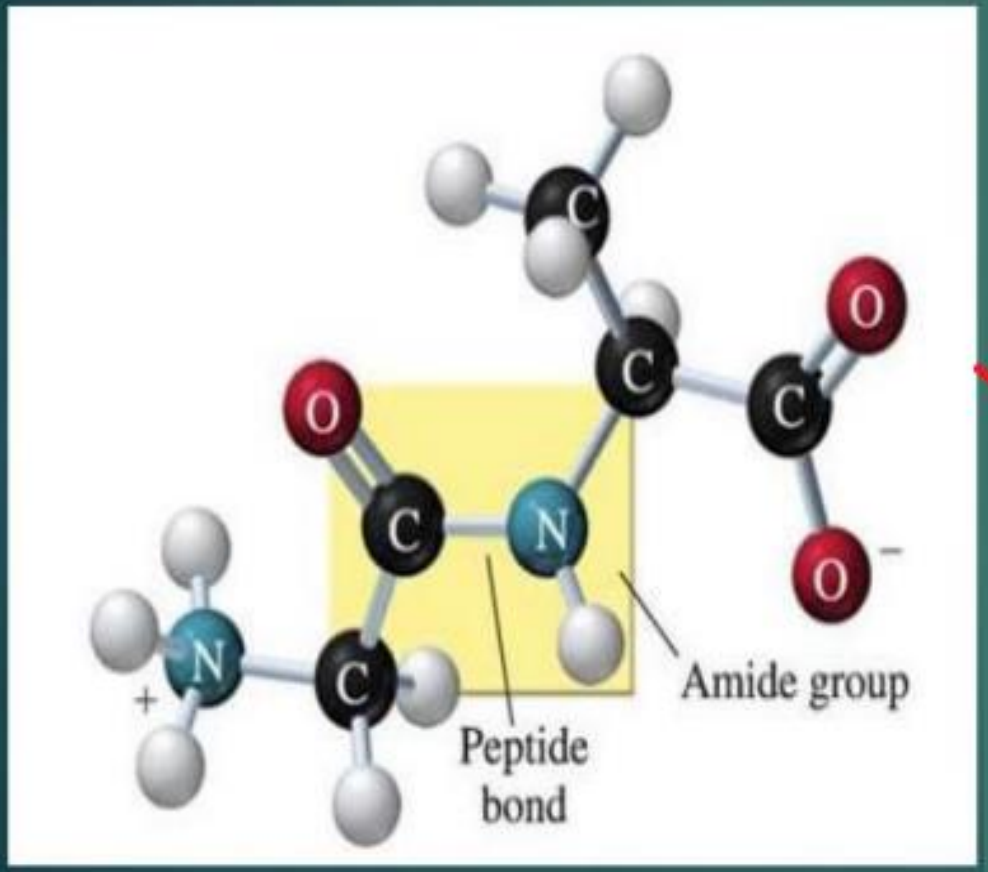
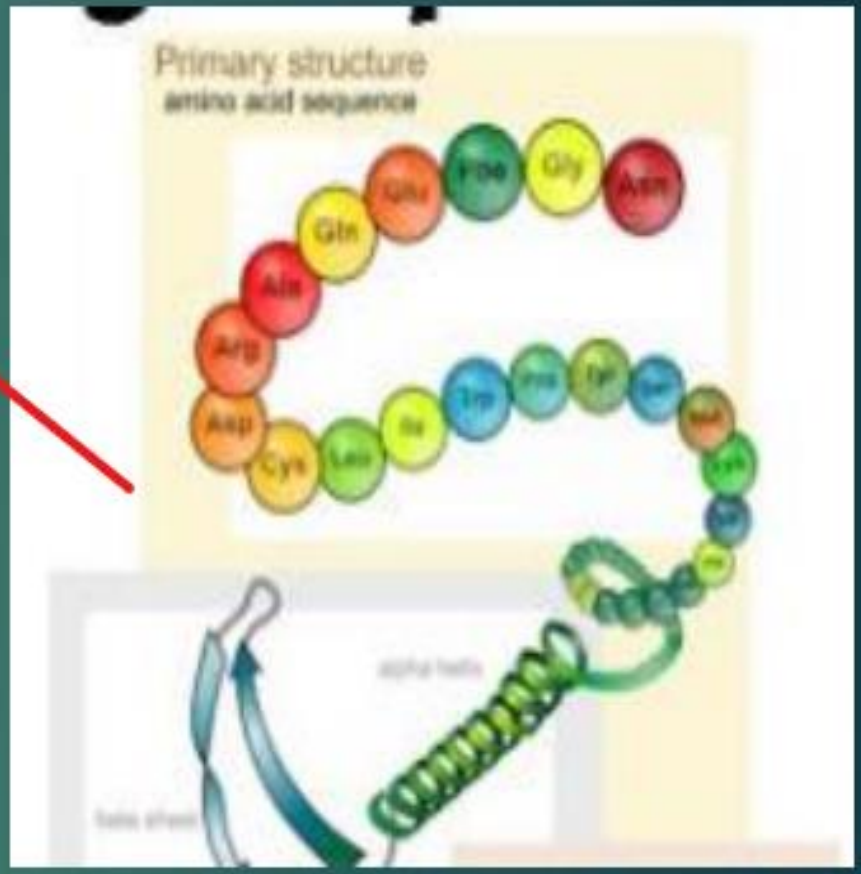
DETERMINATION OF PRIMARY STRUCTURE OF PROTEINS





Determination of primary structure

- The primary structure comprises the identification of amino acids with regard to their quality, quantity and sequence in a protein structure.

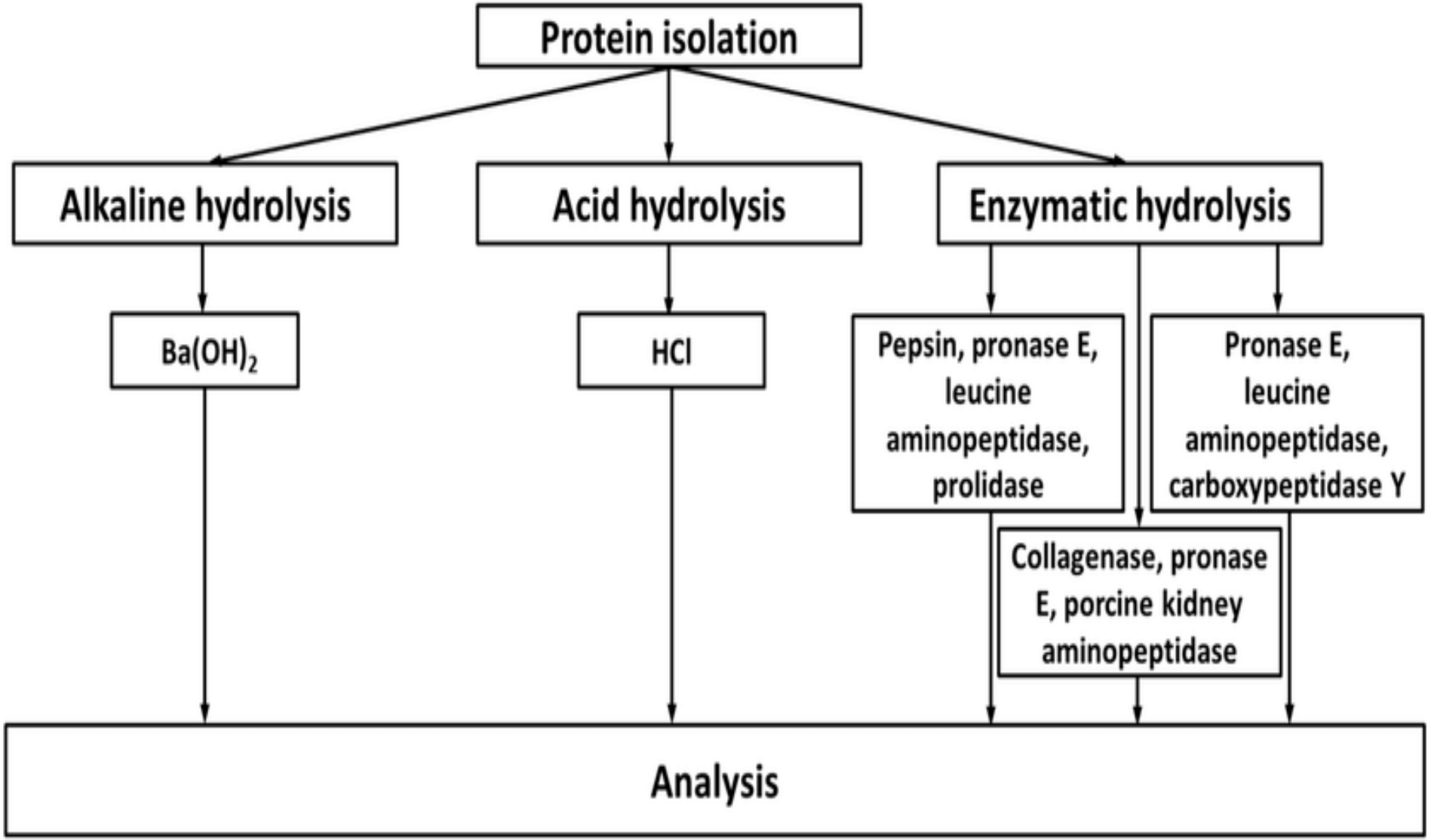
Determination of primary structure involves 3 stages:

- Determination of amino acid composition.
- Degradation of protein or polypeptide into smaller fragments.
- Determination of the amino acid sequence

1. Determination of amino acid composition

- The protein or polypeptide is completely hydrolyzed to liberate the amino acids
- The hydrolysis may be carried out either by acid or alkali treatment or by enzyme hydrolysis.
- *Pronase* is a mixture of non-specific proteolytic enzymes that causes complete hydrolysis of proteins.

Hydrolysis of proteins



I.1

Amino acid composition

Hydrolysis peptide
(HCl 6 M, 110 °C, 24 h)

↓

Seperate amino acids
(HPLC, ion-exchange chromatography, etc)

↓

Analysis results
(comperation of the chromatographic pattern of our sample with that of a standard mixture of amino acids)

↓

Amino acid composition
(elements and concentration)

Ala-Gly-Asp-Phe-Arg-Gly

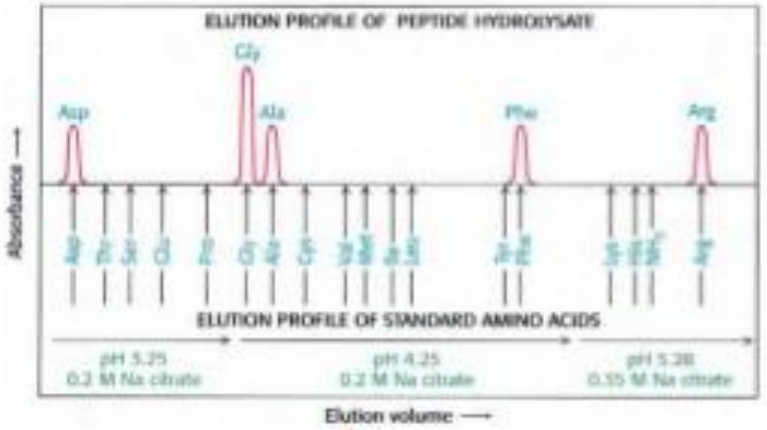
↓

Ala,Gly,Asp,Phe,Arg,Gly

↓

ion-exchange chromatography, ninhydrin reaction

↓



↓

Ala, Arg, Asp, 2Gly, Phe

Steps of protein sequencing

1. Determination of sub-units in polypeptides
2. Determination of number of disulphide bonds (inter and intra-chain)

3. Disruption of tertiary and quaternary structure (Denaturation)
4. Cleavage of polypeptide chain into smaller polypeptides using different chemicals and enzymes
5. Sequencing using Edman's degradation method (specifically identify N-terminal amino acid)
6. Utilization of different enzymes/chemicals for generating several overlapping amino acids and peptides
7. Complete the sequence from overlaps of different peptide frangments in order to determine accurate amino acid sequence in peptides and polypeptides

Determination of number of polypeptides

- Number of polypeptides: The number of polypeptide chains can be identified by treatment of protein with *dansyl chloride.*
- It specifically binds with N-terminal amino acids to form dansyl polypeptides which on hydrolysis yield N-terminal dansyl amino acid.
- The number of dansyl amino acids produced is equal to the number of polypeptide chains in a protein.

Denaturation of Protein

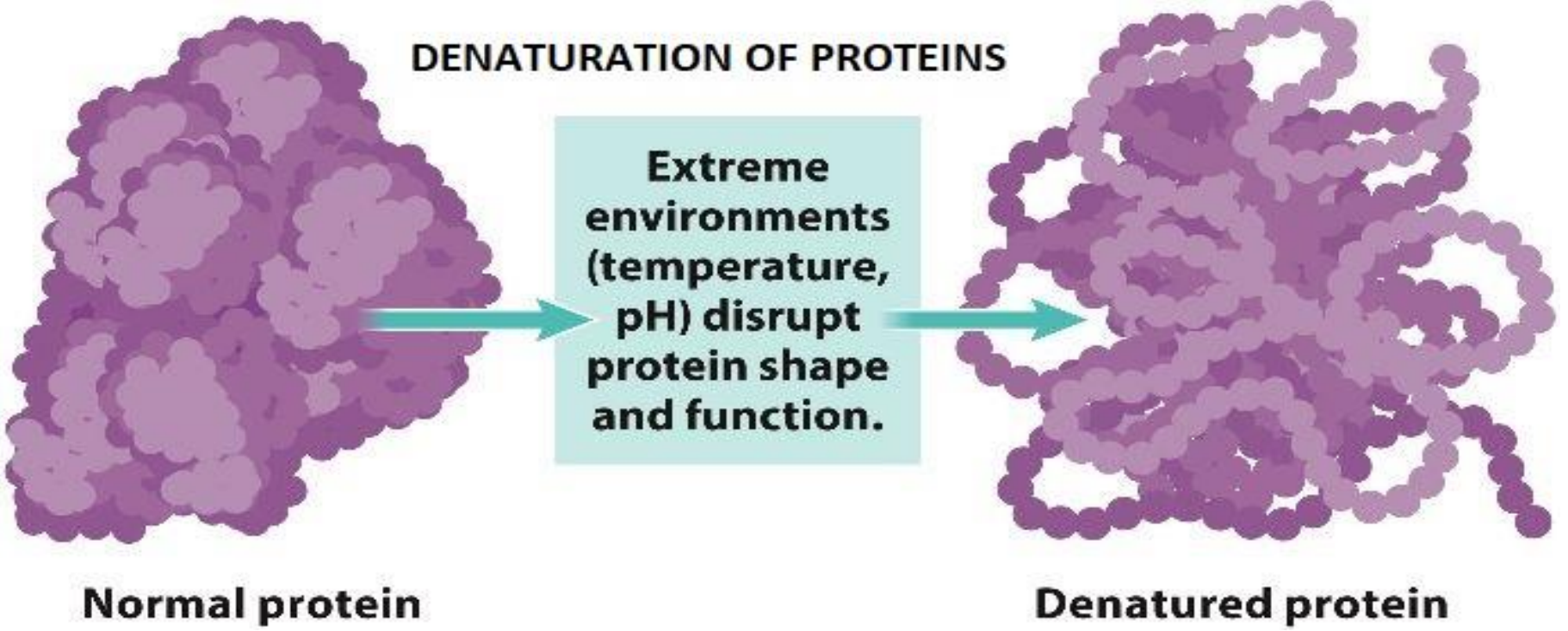
-The term denaturation denotes a reversible or irreversible change of native conformation (tertiary structure) without cleavage of covalent bonds (except for disulfide bridges).

- The primary structure of the protein is not changed because the peptide bonds are not affected

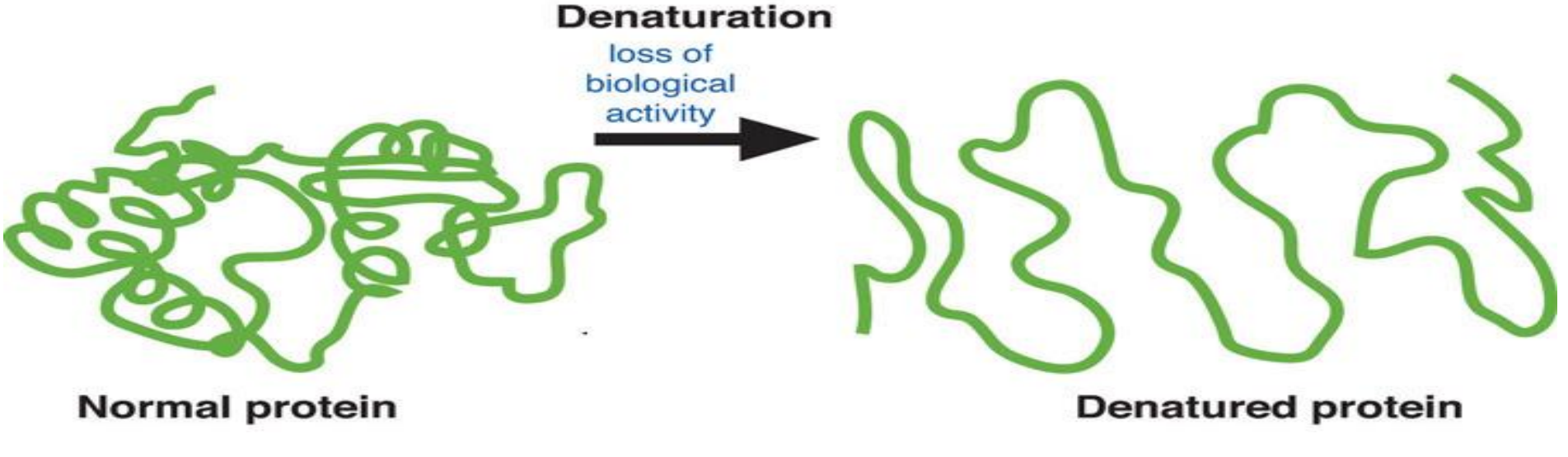
Denaturing agents include:

1. Heat
2. Changes in pH (concentrated acids or alkali)
3. Ultraviolet rays
4. X ray
5. High salt concentration Ex: Detergents- Sodium dodecyl sulphate (SDS)
6. Heavy metals. Ex: Hg^{2+}, Pb^{2+}

7. Urea (6M)- disrupts the hydrogen bonds
8. Reducing agents: Performic acid, 2-Mercaptoethanol, Dithiothereitol (DTT)
   - Breaks disulphide bonds (–S-S- bonds)



agents: pH, temp, ionic strength, solubility



Anfinsen experiment

Christian Anfinsen

- In 1957 Dr. Christian Anfinsen was the first to put protein renaturation on a quantitative basis with the use of bovine pancreatic RNase.
- He was awarded the Nobel Prize in Chemistry in 1972.

- RNase A, a **124-residue** single chain protein, is completely unfolded and its four disulphide bonds reductively cleaved in an **8M urea** solution containing **2-mercaptoethanol.**

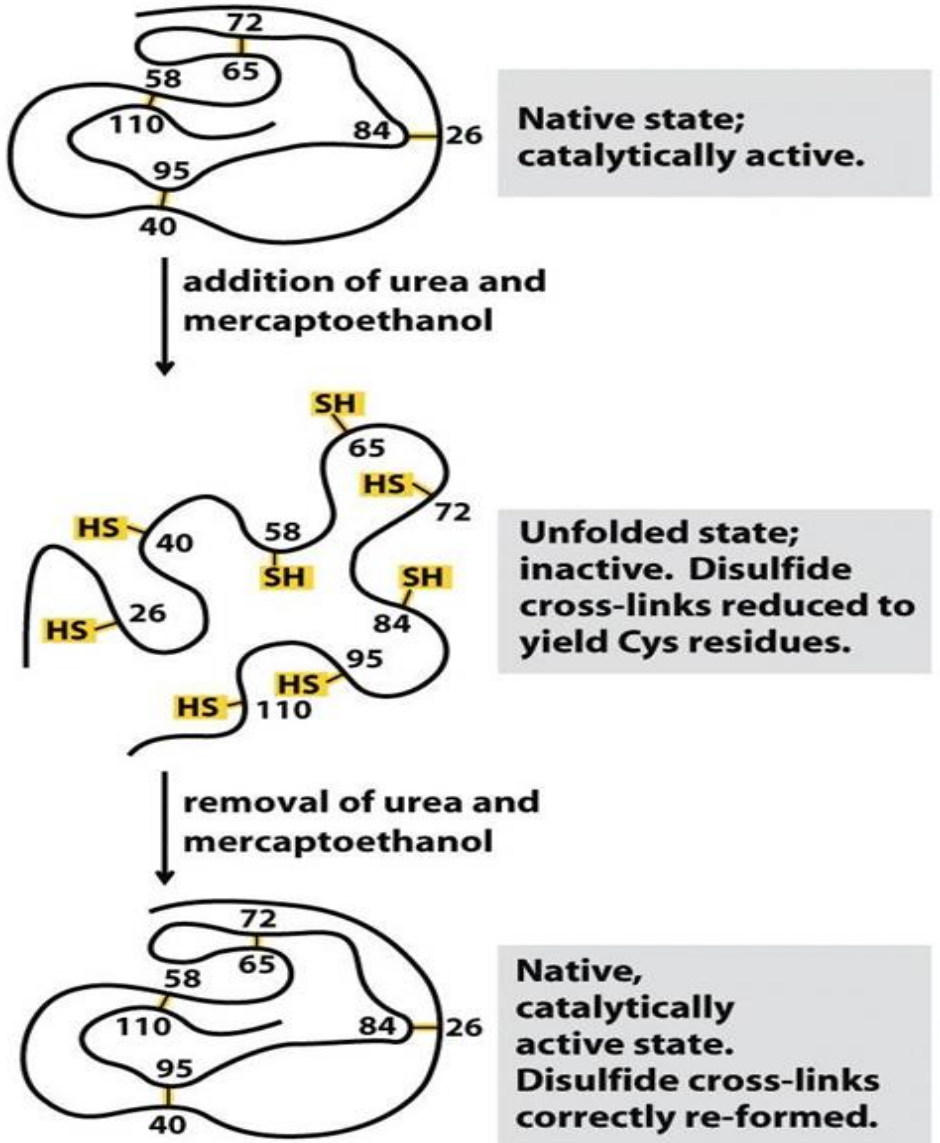


Figure 4-26
Lehninger Principles of Biochemistry, Fifth Edition

Reduction of protein disulfides by β-mercaptoethanol

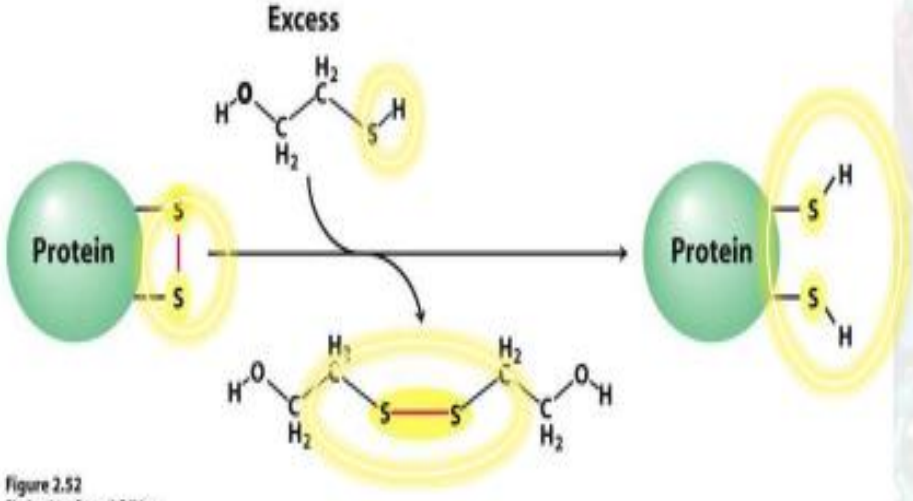
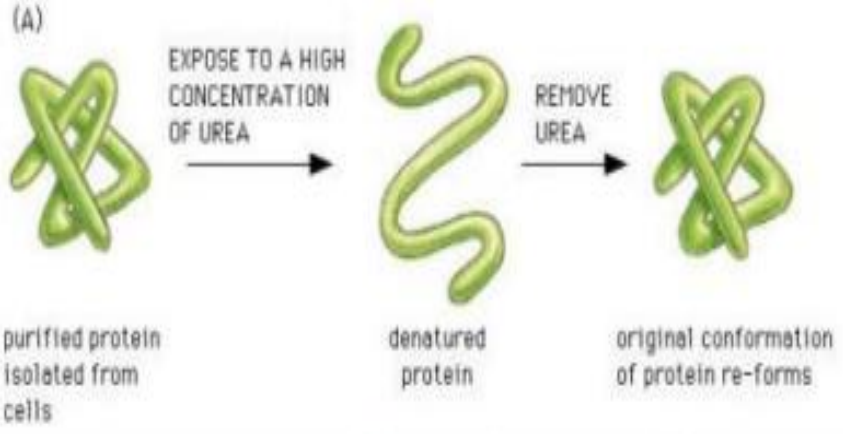


Figure 2.52
Biochemistry, Seventh Edition
© 2012 W. H. Freeman and Company



Proteins - Biomolecules Chemistry Notes | EduRev

2-MERCAPTOETHANOL

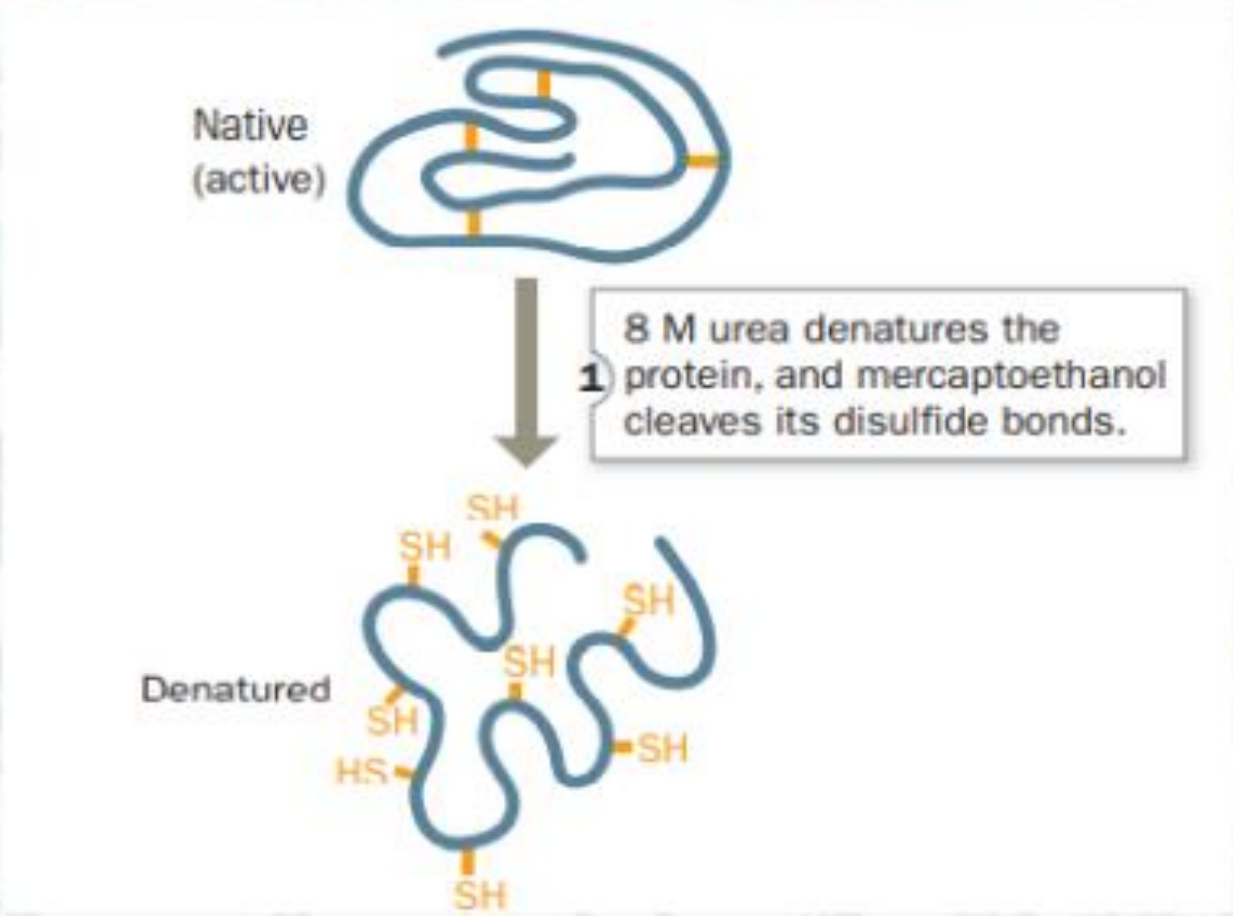
2-mercaptoethanol cleaves the disulfide bonds that may form between thiol groups of cysteine residues.

Urea

Urea interacts with the polar regions of the proteins resulting in the unfolding of the tertiary and secondary structure of the proteins and exposure of the hydrophobic core to water and urea, simply put- urea denatures proteins

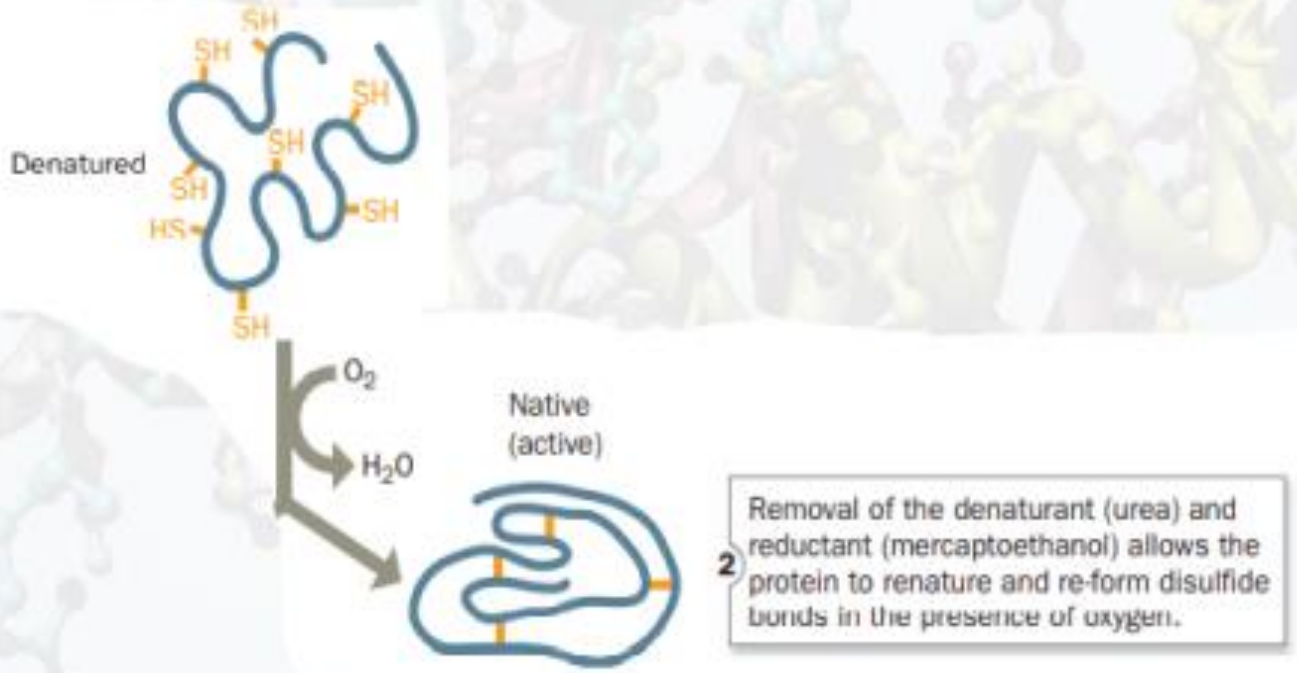
Anfinsen's experiment

RNase A was treated with 8 M urea and 2-mercaptoethanol which resulted in the complete unfolding of the enzyme and the reductive cleavage of all the 4 disulfide bonds



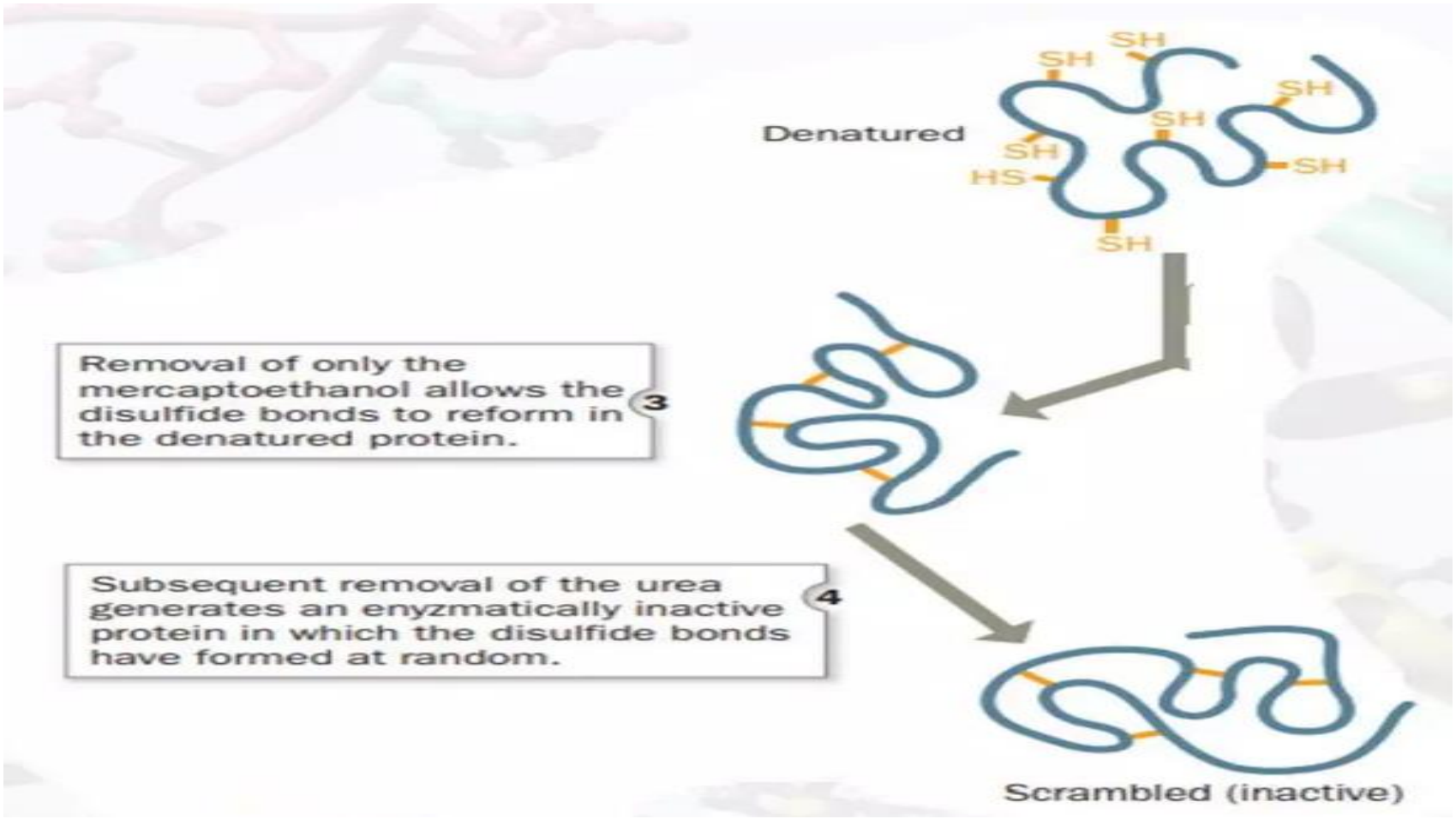
Fundamentals of Biochemistry Life at the Molecular Level by Donald Voet, Charlotte W. Pratt, Judith G. Voet (z-lib.org).pdf

Removal of both urea and 2-mercaptoethanol and exposure to excess oxygen resulted in reformation of the disulfide linkages resulting in protein with 100% enzymatic activity. The resulting protein was physically indistinguishable from the native enzyme.



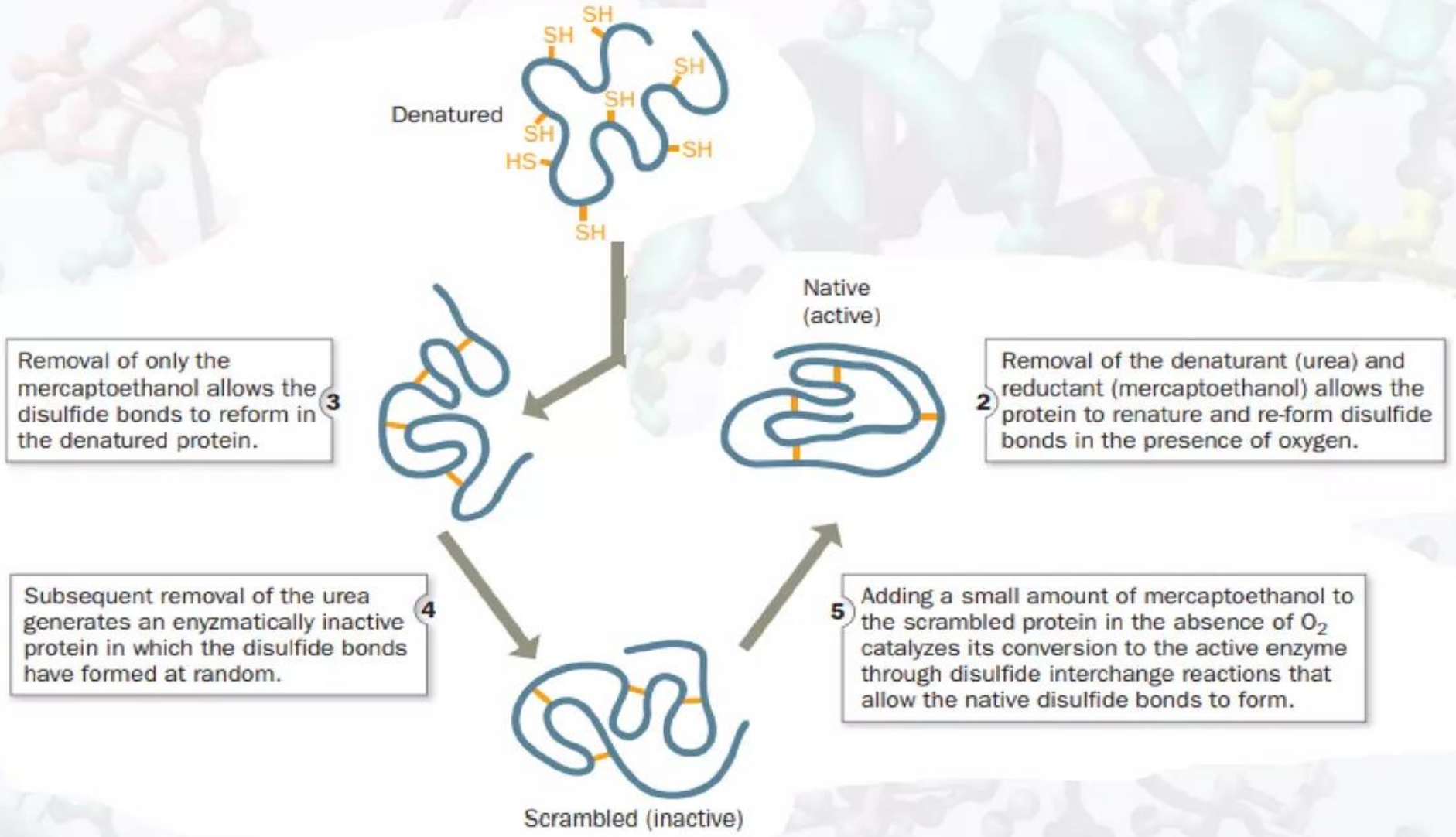
Fundamentals of Biochemistry Life at the Molecular Level by Donald Voet, Charlotte W. Pratt, Judith G. Voet (z-lib.org).pdf

- When RNAse A was oxidized in the presence of urea, protein mixture only has ~1% of the native enzyme activity.
- Indicates, under denaturing conditions disulfide linkage formation is random.
- Removal of urea after the disulfide bonds are formed results in the formation of a scrambled protein.



Fundamentals of Biochemistry Life at the Molecular Level by Donald Voet, Charlotte W. Pratt, Judith G. Voet (z-lib.org).pdf

This “scrambled” protein can be made fully active by exposing it to a **trace of 2-mercaptoethanol**, which breaks the improper disulfide bonds and allows the proper bonds to form.



Fundamentals of Biochemistry Life at the Molecular Level by Donald Voet, Charlotte W. Pratt, Judith G. Voet (z-lib.org).pdf

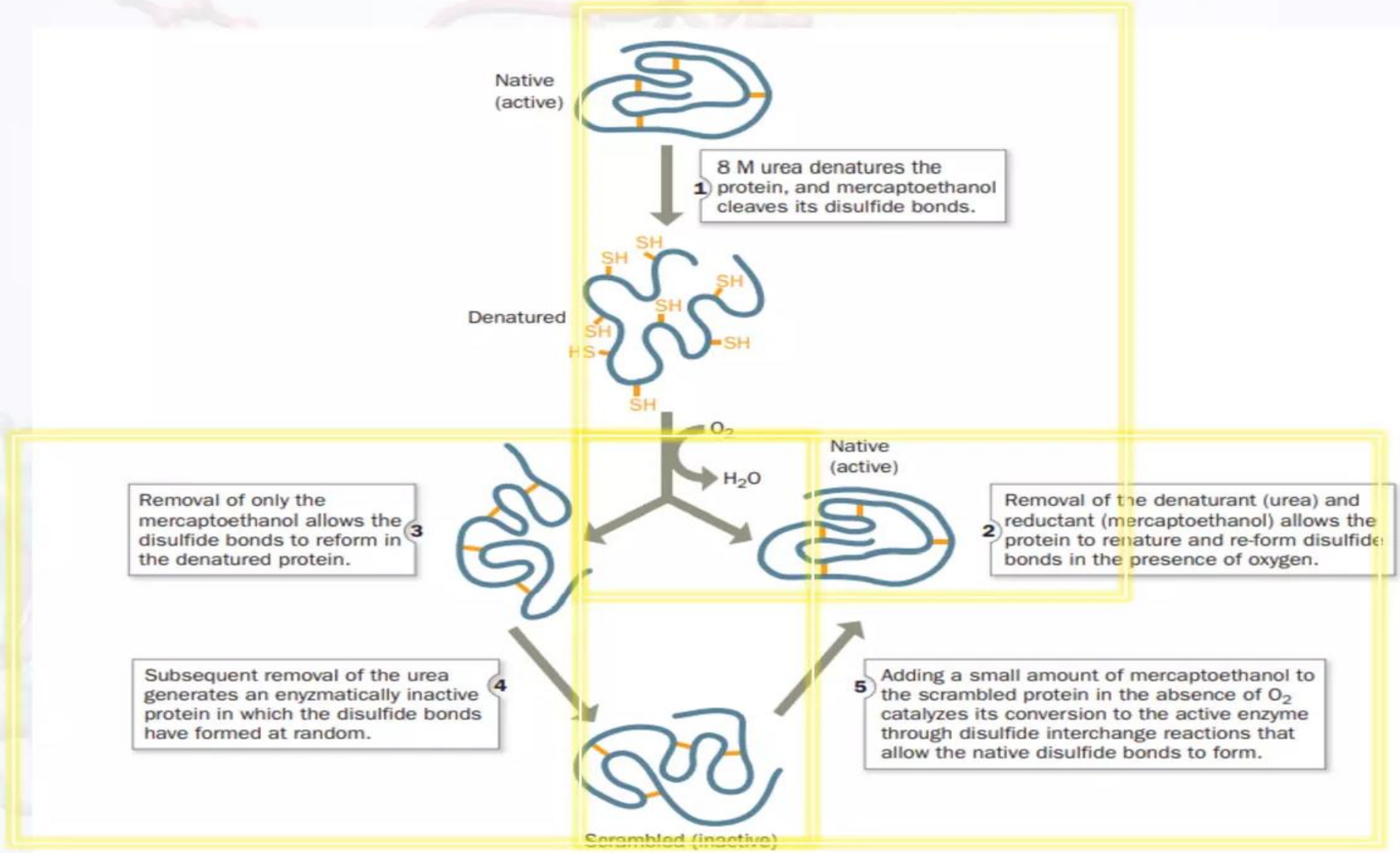


FIG. 6-38 **Denaturation and renaturation of RNase A.** The polypeptide is represented by a blue line, with its disulfide bonds in yellow.

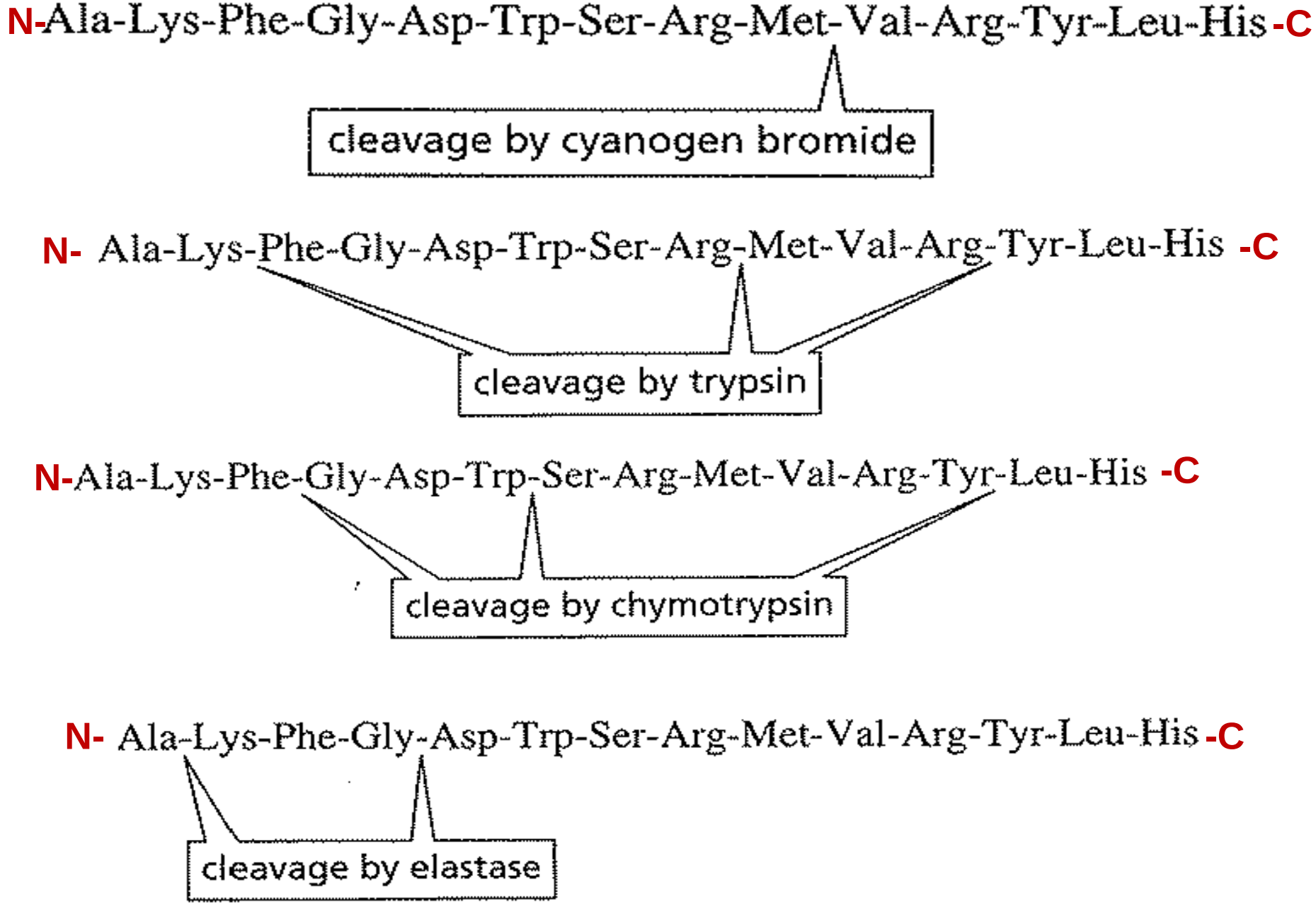
Anfinsen's work demonstrated that proteins can fold spontaneously into their native conformations under physiological conditions. This implies that a protein's primary structure dictates its three dimensional structure.

Breakdown of polypeptides into fragments

- Polypeptides are degraded into smaller peptides by enzymatic or chemical methods
- Enzymatic cleavage: The proteolytic enzymes such as *trypsin, chymotrypsin, pepsin and elastage* exhibit specificity in cleaving the peptide bonds.
- Among these enzymes, trypsin is commonly used.

After denaturation of proteins, various types of chemical treatments or enzymatic treatment can be used to determine the order/sequence of amino acids in protein.

Classification	Reagent	Specificity
Chemical Reagents	Edman's Reagent or phenylisothiocyanate	**Removes N-terminal AA**
	Cyanogen bromide	**Hydrolyzes C-side of Met**
Use of Enzymes	Specific Carboxypeptidase A	**Removes C-terminal AA (NOT Arg or Lys) A for aliphatic and aromatic AA**
	Specific Carboxypeptidase B	**Removes C-terminal AA (ONLY Arg or Lys) B for basic AA**
	Carboxypeptidase (nonspecific)	**Removes C-terminal AA**
	Trypsin	**Hydrolyzes C-side of Arg & Lys**
	Chymotrypsin	**Hydrolyzes C-side of aromatic AA (Phe, Tyr, Trp)**
	Elastase	**Hydrolyzes C-side of small AAs (Gly, Ala)**


N-Ala-Lys-Phe-Gly-Asp-Trp-Ser-Arg-Met-Val-Arg-Tyr-Leu-His-C
cleavage by cyanogen bromide
N- Ala-Lys-Phe-Gly-Asp-Trp-Ser-Arg-Met-Val-Arg-Tyr-Leu-His -C
cleavage by trypsin
N-Ala-Lys-Phe-Gly-Asp-Trp-Ser-Arg-Met-Val-Arg-Tyr-Leu-His -C
cleavage by chymotrypsin
N- Ala-Lys-Phe-Gly-Asp-Trp-Ser-Arg-Met-Val-Arg-Tyr-Leu-His -C
cleavage by elastase

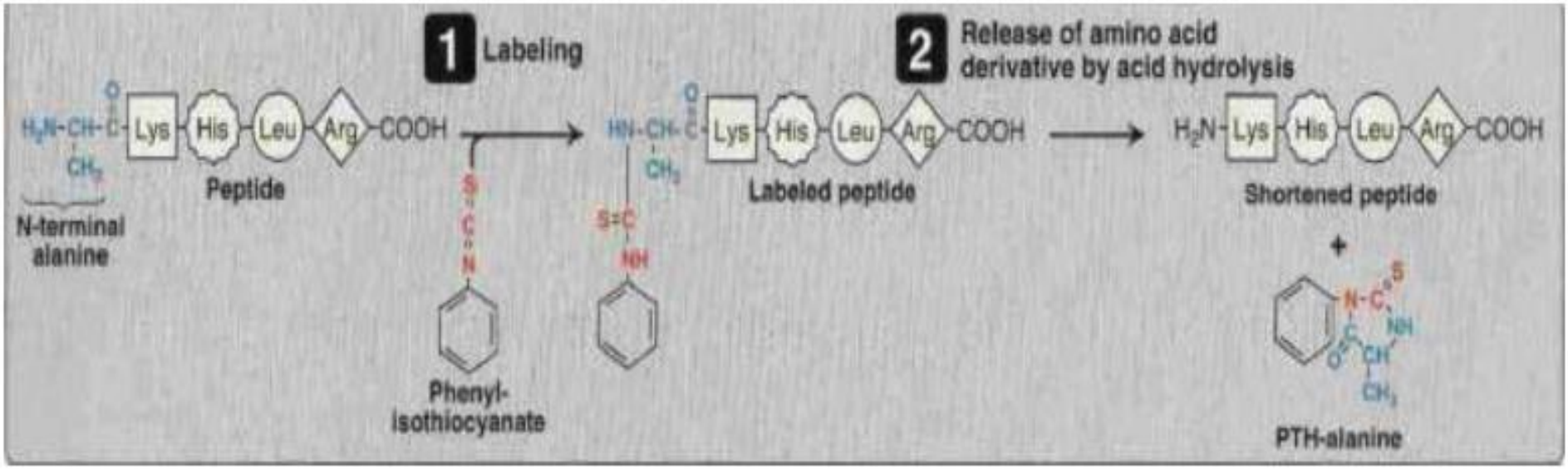
Determination of the amino acid sequence

- Sangers reagent: 1-fluoro-2,4-dinitrobenzene (FDNB)
- FDNB specifically binds with N-terminal amino acids to form a Dinitrophenyl (DNP) derivative of peptide
- This is on hydrolysis yields DNP - amino acids (N-terminal) and free amino acids from the rest of the peptide chain.
- Used for identification of N terminal AA
- DNP-AA - identified by Chromatography

Edmans degradation for amino acid sequencing

- Edmans reagent: Phenyl isothiocyanate is Edmans reagent.
- It reacts with the N-terminal amino acid of peptide to form a phenyl thiocarbamyl derivativve.
- On treatment with mild acid, phenyl thiohydrantoin (PTH) –amino acid, a cyclic compound is liberated.

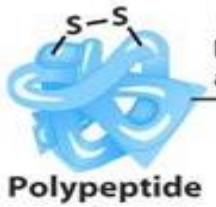
- PTH amino acid can be identified by chromatography
- Edmans reagent has an advantage, a peptide can be sequentially degraded liberating N-terminal amino acids one after another which can be identified.
- This is due to the fact that the peptide as a whole is not hydrolyzed but only releases PTH- amino acids.



Overlapping peptides

- In the determination of primary structure of protein, several methods are simultaneously employed.
- This results in the formation of overlapping peptides
- Overlapping peptides are very useful in determining the amino acid sequence.

Protein Sequencing: Overlapping Sequences



Polypeptide

Procedure

hydrolyze; separate amino acids

Result

A	5	H	2	R	1
C	2	I	3	S	2
D	4	K	2	T	1
E	2	L	2	V	1
F	1	M	2	Y	2
G	3	P	3		

Conclusion

Polypeptide has 38 amino acid residues. Trypsin will cleave three times (at one R (Arg) and two K (Lys)) to give four fragments. Cyanogen bromide will cleave at two M (Met) to give three fragments.

react with FDNB; hydrolyze; separate amino acids

2,4-Dinitrophenylglutamate detected

E (Glu) is amino-terminal residue.

reduce disulfide bonds (if present)

cleave with trypsin; separate fragments; sequence by Edman degradation

- T-1 GASMALIK
- T-2 EGAAYHDFEPIDPR
- T-3 DCVHSD
- T-4 YLIACGPMTK

T-2 placed at amino terminus because it begins with E (Glu).

T-3 placed at carboxyl terminus because it does not end with R (Arg) or K (Lys).

cleave with cyanogen bromide; separate fragments; sequence by Edman degradation

- C-1 EGAAYHDFEPIDPRGASM
- C-2 TKDCVHSD
- C-3 ALIKYLIACGPM

C-3 overlaps with T-1 and T-4, allowing them to be ordered.

establish sequence

Amino terminus

Carboxyl terminus

Sequenator

- This is an automatic machine to determine the amino acid sequence in a polypeptide
- It is based on the principle of Edman's degradation .
- Amino acids are determined sequentially from N-terminal end
- The PTH-amino acid liberated is identified by HPLC.
- Sequenator takes about 2 hours to determine each amino acid.

Reverse sequencing technique

- It is the genetic material (DNA) which ultimately determines the sequence of amino acids in a polypeptide chain
- By analyzing the nucleotide sequence of DNA that codes for protein, it is possible to translate the nucleotide sequence into amino acid sequence.
- This technique, however, fails to identify the disulfide bonds and changes that occur in the amino acids after the protein is synthesized.

• Steps for Determining the Primary Structure

Steps	Technique
I. Determination of amino acid composition in a protein	
Complete hydrolysis of protein	Pronase enzyme
Separation and estimation of amino acids	Chromatographic technique
II. Degradation of protein into smaller fragments	
Liberation of polypeptides	Urea or guanidine hydrochloride
Determination of the number of polypeptide chains in a protein	Dansyl chloride
Breakdown of polypeptides into fragments	Enzymatic cleavage Chemical cleavage
III. Determination of amino acid sequence	Sanger's reagent Edman's reagent
Nowadays, DNA sequencing is used to determine the amino acid sequence.	

Sample questions

1. Which of the following information is responsible to specify the three-dimensional shape of a protein?

a) Protein's peptide bond
b) Protein's amino acid sequence
c) Protein's interaction with other polypeptides
d) Protein's interaction with molecular chaperons

2. A new drug is developed which selectively cleaves covalent bonds between two sulphur atoms of non-adjacent amino acids in a polypeptide chain. Which level of protein structure in affected molecules would be most directly affected by the drug?

a) Primary structure
b) Secondary structure
c) Tertiary structure
d) Quaternary structure

3. State TRUE or FALSE

I. Quarternary structure refers to the three-dimensional arrangement of polypeptide subunits in a protein consisting of two or more polypeptide chains.
II. The cis isomer is the highly favourable isomer for most peptide bonds except those preceding proline.
III. Ramachandran plot is a graphical representation of the sterically allowed conformations of peptide planes

4. Which structure of a protein is the arrangement of protein subunits in a multi-subunit complex?

a) Primary
b) Secondary
c) Tertiary
d) Quaternary

5. Which of the following options contain only non-covalent interactions?

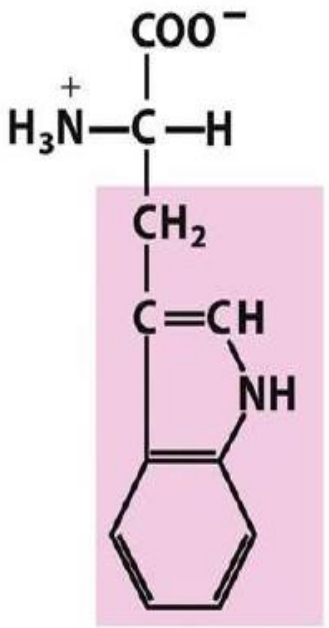
a) Disulfide bond, Ionic bond
b) Ionic bond, peptide bond
c) Hydrogen bond, disulfide bond
d) Hydrogen bond, Ionic bond

6. Albumin is a large protein which circulates freely in human plasma. Another type of protein, called aquaporins, act as a channel for water to enter and exit a cell, and are located within the lipid bilayer of a cell's plasma membrane. Based on their physiological locations, how would the tertiary structure of these two proteins mostly likely compare?

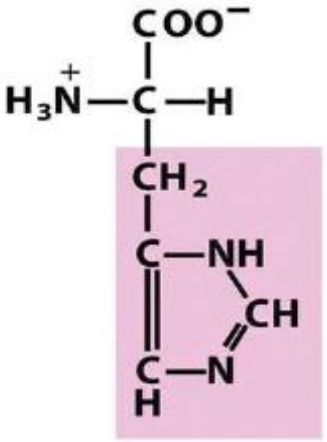
(A) Albumin has a more nonpolar core with a nonpolar outer layer, while aquaporins have a more polar core with a polar outer layer
(B) Albumin has a more nonpolar core with a polar outer layer, while aquaporins have a more polar core with a nonpolar outer layer
(C) Albumin has a more polar core with a nonpolar outer layer, while aquaporins have a more nonpolar core with a polar outer layer
(D) Albumin has a more polar core with a polar outer layer, while aquaporins have a more nonpolar core with a nonpolar outer layer

(7) For the following five amino acids, examine their side chains and determine whether the side chain is nonpolar, acidic, or basic. Assume each amino acid is at physiological pH.

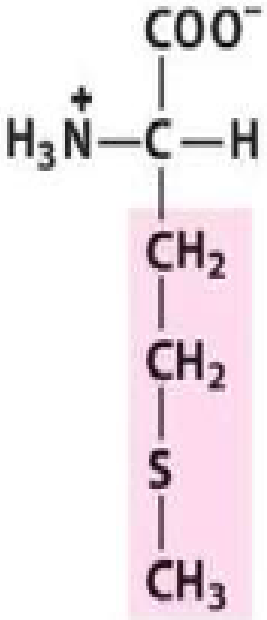
- Tryptophan
- Histidine
- Methionine
- Phenylalanine
- Glutamic acid



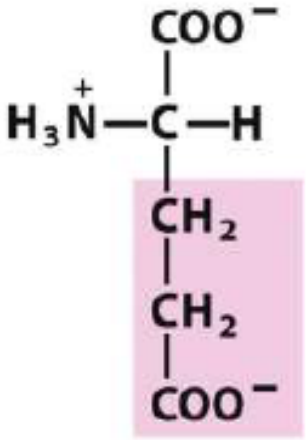
Tryptophan



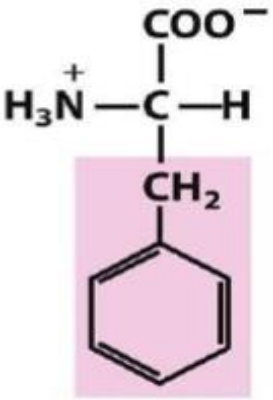
Histidine



Methionine



Glutamate



Phenylalanine

(8) For each of the following functional groups, name at least one amino acid that has one in its structure.

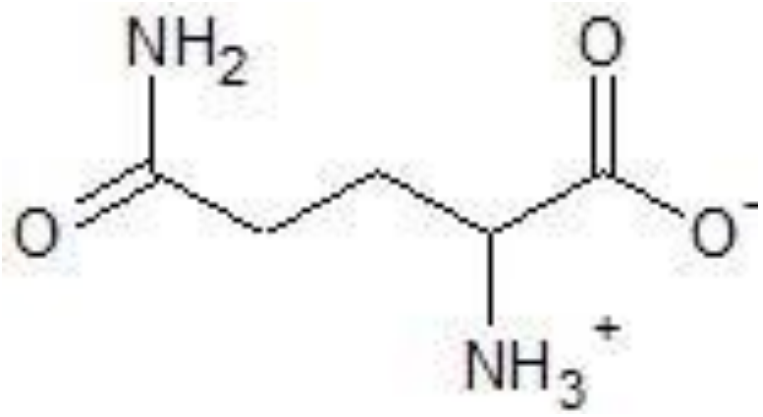
- An aromatic ring
- An alcohol
- An amide
- A phenol
- A sulfhydryl group

(9) Name the four different types of interactions responsible for creating tertiary structures in proteins.

(10) Name at least three amino acids whose side chains have the characteristic of being able to repel water.

(11) Define denaturation of proteins and list five different ways to denature a protein?

(9) An example of an amino acid is shown below. Answer the following questions relating to its structure.



- Circle the amine functional group. Is the amine neutral or ionic?
- Put a triangle around the carboxylate functional group. What is the one thing that differs between a carboxylate and a carboxylic acid?
- What is the definition of a zwitterion? Does this particular amino acid fit the definition? Name the amino acids that can NOT be considered a zwitterion at physiological pH.
- Box the entire side chain of the amino acid. Use this to identify the name of this amino acid.
- Describe the character of the side chain. Is it nonpolar, polar, acidic, or basic?
- Sketch the amino acid and what it would look like if you raised the pH up to 11. Is it positively charged, negatively charged or neutral overall? Repeat this with the pH changed down to 4.

Sample questions

Question 1: A pentapeptide was found to have the composition Ala Arg Gly Pro Trp. Reaction of the pentapeptide with Sanger's reagent, followed by hydrolysis, gave the DNP derivative of proline.

- **Treatment of the pentapeptide with carboxypeptidase initially produced alanine.**

- **Treatment of the pentapeptide with trypsin gave a tetrapeptide which, when treated with chymotrypsin, produced a tripeptide.**

Which one of the following is the sequence of the pentapeptide? (note: trypsin cleaves after basic amino acids)

A) Pro-Gly-Arg-Trp-Ala
B) Pro-Arg-Gly-Trp- Ala
C) Pro-Trp-Arg-Gly-Ala
D) Pro-Gly-Trp-Arg-Ala
E) Pro-Trp-Gly-Arg-Ala

Question 2: Oxytocin, a hormone peptide of nine amino acids is widely used in obstetrics to induce uterine contractions. There is an intramolecular disulfide bond which must be reduced before sequencing. Reduced oxytocin has the composition Asn Cys2 Gln Gly Ile Leu Pro Tyr. Partial hydrolysis of reduced oxytocin led to the following fragments.

Asn-Cys
Cys-Tyr
Tyr-Ile-Gln
Cys-Pro-Leu
Ile-Gln
Leu-Gly
Gln-Asn-Cys

- Reaction of reduced oxytocin with carboxypeptidase showed glycine as the first liberated amino acid. Which one of the following is the sequence of oxytocin?

A) Asn Cys Tyr Ile Gln Cys Pro Leu Gly
B) Gln Asn Cys Tyr Ile Cys Pro Leu Gly
C) Cys Pro Gln Asn Cys Tyr Ile Leu Gly
D) Cys Tyr Ile Gln Asn Cys Pro Leu Gly
E) Gln Asn Cys Tyr Ile Cys Pro Leu Gly

Practice questions on peptide cleavage

3. Given is the peptide sequence of 16 amino acids with peptide bonds in between them.

L-M-A-G-D-S-N-K-V-G-I-R-L-T-W-H…. This peptide is treated with Cyanogen bromide, Trypsin and Chymotrypsin at the same time.

Write the products of cleavage and name them as di-; tri-; tetra-; penta-; hexa- so on…..

Answer: ????

Practice questions on peptide cleavage

4) An octapeptide contains the following amino acids Ile, Arg, Met, Phe, Pro, Thr, Tyr, Val.

- **Trypsin digestion yields a dipeptide containing Thr-Arg, and remaining hexapeptide;**
- **Chymotrypsin digestion yields two tetrapeptides with amino acid composition as Thr-Arg-Val-Tyr and the other containing Ile-Met-Pro-Phe**
- **A chemical digestion with cyanogen bromide yields a dipeptide containing Pro-Phe and remaining amino acids of hexapeptide. Cyanogen bromide cleaves on the C-terminal side of the Methionine.**
- **What will be the sequence of amino acids in octapeptide???**

Answer: ????

Practice questions on peptide cleavage

5. A peptide with 12 amino acids has the composition (not in sequence)

Asp Cys_2 Glu_2 Leu_2 Ser_2 Tyr_2 Val. It also has Ser as the N-terminal residue and Cys as C-terminal. Partial acid hydrolysis gave these peptide sequences.

Ser Leu Tyr
Tyr Cys
Leu Tyr Glu
Glu Leu Glu
Ser Val Cys
Cys Ser Val
Glu Asp Tyr

Do overlap analysis and determine the amino acid sequence of the original peptide.

Practice Questions on peptide cleavage

6. You have to determine the amino acid sequence of a peptide. You perform the following steps using enzyme cleavage of your peptide:

Step 1. Treatment with **trypsin** yields three fragments with the following sequences (in the order of their length): WGA, AGTK, YLDR

Step 2. Treatment with **chymotrypsin** gave the following three peptide fragments: GA, LDRW, AGTKY.

What is the sequence of your peptide ?????