1

Chem 352 - Lecture 3
Part I: Amino Acids and Protein Primary Structure

Question for the Day: Approximately how much mass is required to make just one molecule for each of the possible polypeptides with a length of 100 amino acid residues, which are made from the 20 naturally occurring L-amino acids?

2

Introduction to Proteins

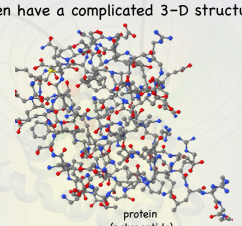
Proteins are the workhorses of a living cell.

- Biological catalysts (enzymes)
- Storage and transport
- Cytoskeleton
- Cellular regulation
- Hormones
- Antibodies
- ...

3

Introduction to Proteins

Proteins are polymers of amino acids (polypeptides) that often have a complicated 3-D structures.



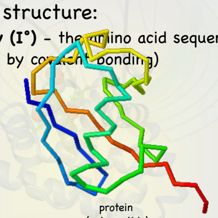
protein
(polypeptide)

4

Introduction to Proteins

Proteins can have up to four different levels of structure:

- **Primary (I°)** - the amino acid sequence (defined by covalent bonding)



protein
(polypeptide)

5-1

Introduction to Proteins

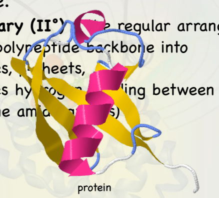
Proteins have different levels of structure:

- **Secondary (II°)** - the regular arrangements of the polypeptide backbone into α-helices, β-sheets, ...
(features hydrogen bonding between the backbone amide groups)

Introduction to Proteins

Proteins have different levels of structure:

- **Secondary (II°)** – the regular arrangements of the polypeptide backbone into α-helices, β-sheets, … (features hydrogen bonding between the backbone amide groups)



protein (polypeptide)

Chem 352, Lecture 3 – Amino Acids and Protein Primary Structure 5

5-2

Introduction to Proteins

Proteins have different levels of structure:

- **Secondary (II°)** – the regular arrangements of the polypeptide backbone into α-helices, β-sheets, … (features hydrogen bonding between the backbone amide groups)

Chem 352, Lecture 3 – Amino Acids and Protein Primary Structure 5

5-3

Introduction to Proteins

Proteins have different levels of structure:

- **Tertiary (III°)** – the 3-dimensional fold of a single polypeptide (stabilized by non-covalent side chain interactions)

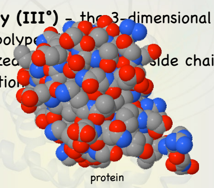
Chem 352, Lecture 3 – Amino Acids and Protein Primary Structure 6

6-1

Introduction to Proteins

Proteins have different levels of structure:

- **Tertiary (III°)** – the 3-dimensional fold of a single polypeptide (stabilized by non-covalent side chain interactions)



protein (polypeptide)

Chem 352, Lecture 3 – Amino Acids and Protein Primary Structure 6

6-2

Introduction to Proteins

Proteins have different levels of structure:

- **Tertiary (III°)** – the 3-dimensional fold of a single polypeptide (stabilized by non-covalent side chain interactions)

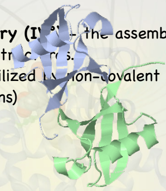
Chem 352, Lecture 3 – Amino Acids and Protein Primary Structure 6

6-3

Introduction to Proteins

Proteins have different levels of structure:

- **Quaternary (IV°)** - the assembly of multiple tertiary structures. (also stabilized by non-covalent side chain interactions)

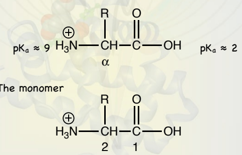


protein with two subunits (2 polypeptides)

7

The Amino Acids

- Proteins are made from polymers of α-amino acids.
  - These polymers are called **polypeptides**.
  - The **monomers** are the α-amino acids.

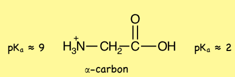


8-1

The Amino Acids

Problem:

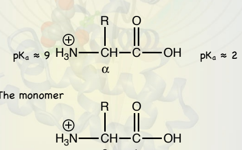
Draw the titration curve (pH vs. equivalents) for glycine between pH 0 and pH 12.



8-2

The Amino Acids

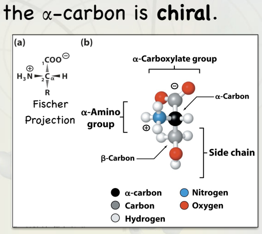
- Proteins are made from polymers of α-amino acids.
  - These polymers are called **polypeptides**.
  - The **monomers** are the α-amino acids.



8-3

The Amino Acids

For 19 out of the 20 common amino acids, the α-carbon is **chiral**.



9-1

9-2

The Amino Acids

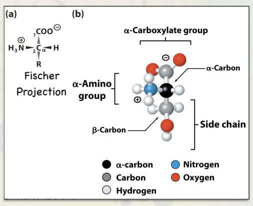
Question:

What does it mean to be **chiral**?

9-3

The Amino Acids

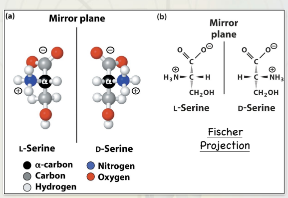
For 19 out of the 20 common amino acids, the α-carbon is **chiral**.



10-1

The Amino Acids

For 19 out of the 20 common amino acids, the α-carbon is **chiral**.



10-2

The Amino Acids

Question:

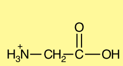
Which of the 20 amino acids is not **chiral**?

10-3

The Amino Acids

Question:

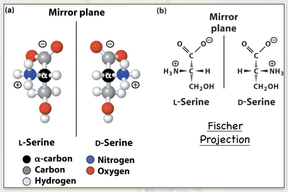
Which of the 20 amino acids is not **chiral**?



Glycine

The Amino Acids

For 19 out of the 20 common amino acids, the α-carbon is **chiral**.



10-4

The Amino Acids

There are different ways to designate the stereochemistry of a chiral center:

11-1

The Amino Acids

There are different ways to designate the stereochemistry of a chiral center:

- **R vs S**
  - Rectus versus Sinister - based on the atomic mass of the substituents

11-2

The Amino Acids

There are different ways to designate the stereochemistry of a chiral center:

- **R vs S**
  - Rectus versus Sinister - based on the atomic mass of the substituents
- **d (+) vs l (-)**
  - dextrorotatory versus levorotatory - based on the bending of plane polarized light

11-3

The Amino Acids

There are different ways to designate the stereochemistry of a chiral center:

- **R vs S**
  - Rectus versus Sinister - based on the atomic mass of the substituents
- **d (+) vs l (-)**
  - dextrorotatory versus levorotatory - based on the bending of plane polarized light
- **D vs L**
  - Based on how glyceraldehyde bends plane polarized light

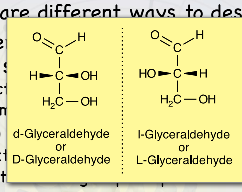
11-4

11-5

The Amino Acids

There are different ways to designate the ste... center:

- R vs S
  - Rect... he atom...
- d (+)
  - dext... based on t... ght
- D vs L
  - Based on how glyceraldehyde bends plane polarized light



11-6

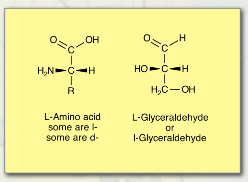
The Amino Acids

There are different ways to designate the stereochemistry of a chiral center:

- **R vs S**
  - Rectus versus Sinister - based on the atomic mass of the substituents
- **d (+) vs l (-)**
  - dextrorotatory versus levorotatory - based on the bending of plane polarized light
- **D vs L**
  - Based on how glyceraldehyde bends plane polarized light

12-1

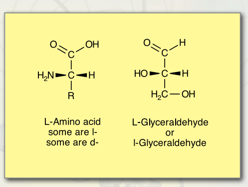
The Amino Acids



12-2

The Amino Acids

All of the amino acids used to to make proteins are L-amino acids



13

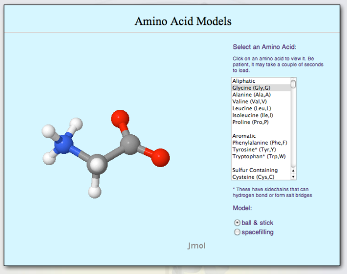
The Amino Acids

There is considerable variety in the chemical and physical properties of the the 20 different amino acid side chains.

- Aliphatic (saturated hydrocarbon) (G,A,V,L,I,P)
- Aromatic (F,Y,W)
- Sulfur-containing (C,M)
- Alcohols (S,T)
- Bases (K,R,H)
- Acids (D,E)
- Amides (N,Q)

14

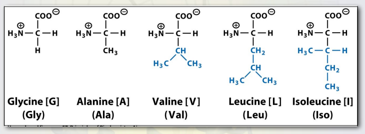
The Amino Acids



15

The Amino Acids

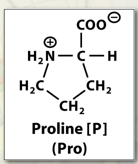
- Aliphatic (saturated hydrocarbon) (G,A,V,L,I,P)



16

The Amino Acids

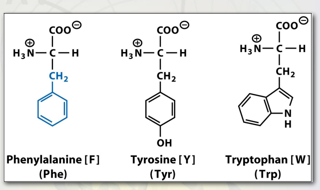
- Aliphatic (saturated hydrocarbon) (G,A,V,L,I,P)



17-1

The Amino Acids

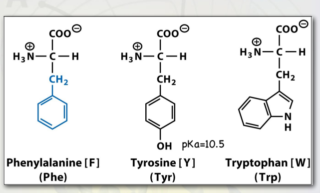
- Aromatic (F,Y,W)



17-2

The Amino Acids

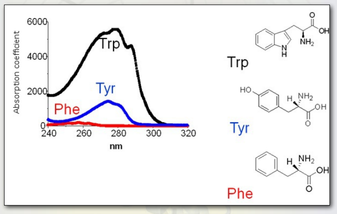
- Aromatic (F,Y,W)



17-3

The Amino Acids

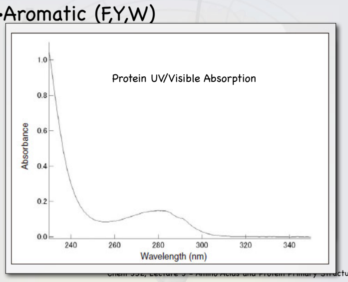
- Aromatic (F,Y,W)



17-4

The Amino Acids

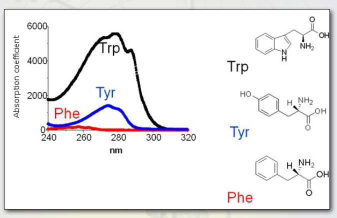
- Aromatic (F,Y,W)



17-5

The Amino Acids

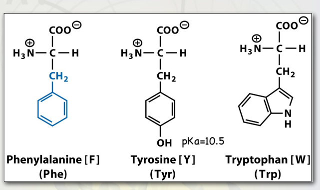
- Aromatic (F,Y,W)



17-6

The Amino Acids

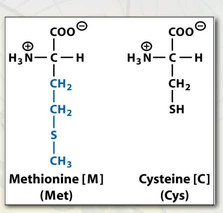
- Aromatic (F,Y,W)



18-1

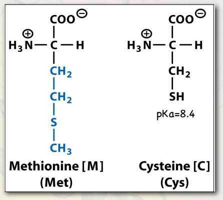
The Amino Acids

- Sulfur-containing (C,M)



The Amino Acids

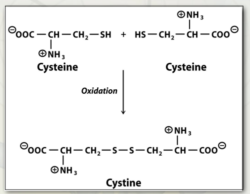
- Sulfur-containing (C,M)



18-2

The Amino Acids

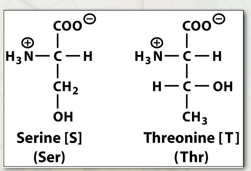
- Sulfur-containing (C,M)



19

The Amino Acids

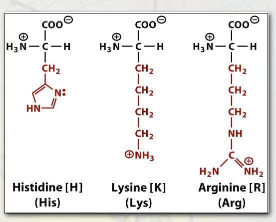
- Alcohols (S,T)



20

The Amino Acids

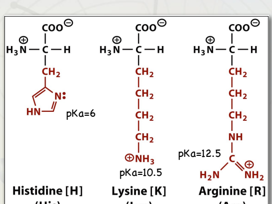
- Bases (H,K,R)



21-1

The Amino Acids

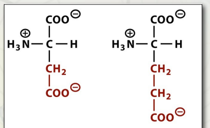
- Bases (H,K,R)



21-2

The Amino Acids

- Acids (D,E)

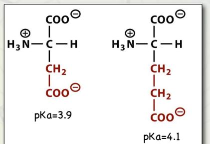


Aspartate [D] (Asp) | Glutamate [E] (Glu)

22-1

The Amino Acids

- Acids (D,E)

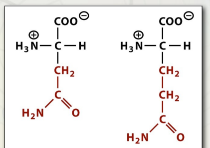


Aspartate [D] (Asp) | Glutamate [E] (Glu)

22-2

The Amino Acids

- Amides (N,Q)



Aspargine [N] (Asn) | Glutamine [Q] (Gln)

23

The Amino Acids

- Can also group the amino acids based on their solubility in water.
  - Highly hydrophobic (I, P, V, L, M)
  - Less hydrophobic (W, A, G, C, Y, P, T, S)
  - Highly hydrophilic (H, E, N, Q, D, K, R)

24-1

The Amino Acids

- Can also ... acids based on their
  - Highly h...)
  - Less hy... P, T, S)
  - Highly h... K, R)

TABLE 3.1 Hydropathy scale for amino acid residues

Amino acid	Free-energy change for transfer[a] (kj mol^{-1})
Highly hydrophobic	
Isoleucine	3.1
Phenylalanine	2.5
Valine	2.3
Leucine	2.2
Methionine	1.1
Less hydrophobic	
Tryptophan	1.5[b]
Alanine	1.0
Glycine	0.67
Cysteine	0.17
Tyrosine	0.08
Proline	−0.29
Threonine	−0.75
Serine	−1.1
Highly hydrophilic	
Histidine	−1.7
Glutamate	−2.6
Asparagine	−2.7
Glutamine	−2.9
Aspartate	−3.0
Lysine	−4.6
Arginine	−7.5

24-2

The Amino Acids

- Can also group the amino acids based on their solubility in water.
  - Highly hydrophobic (I, P, V, L, M)
  - Less hydrophobic (W, A, G, C, Y, P, T, S)
  - Highly hydrophilic (H, E, N, Q, D, K, R)

24-3

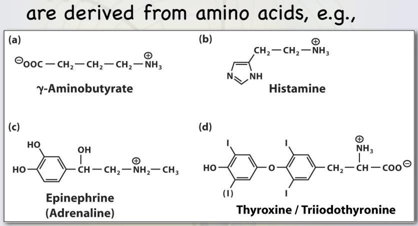
The Amino Acids

- Many important biological molecules are derived from amino acids, e.g.,
  - **histamine** (derived from His) (triggers allergic response)
  - **epinephrine** (derived from Tyr) (hormone that triggers "flight or fight" response)
  - **thyroxine** (derived from Tyr) (thyroid hormone) (one of the few uses for iodine)
  - **GABA** (derived from Glu) (neurotransmitter)

25-1

The Amino Acids

- Many important biological molecules are derived from amino acids, e.g.,



25-2

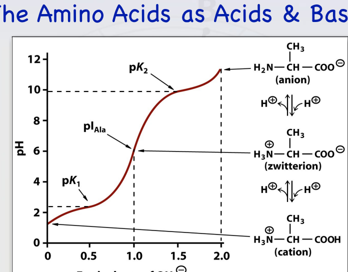
The Amino Acids

- Many important biological molecules are derived from amino acids, e.g.,
  - **histamine** (derived from His) (triggers allergic response)
  - **epinephrine** (derived from Tyr) (hormone that triggers "flight or fight" response)
  - **thyroxine** (derived from Tyr) (thyroid hormone) (one of the few uses for iodine)
  - **GABA** (derived from Glu) (neurotransmitter)

Chem 352, Lecture 3 - Amino Acids and Protein Primary Structure 25

25-3

The Amino Acids as Acids & Bases



26-1

The Amino Acids as Acids & Bases

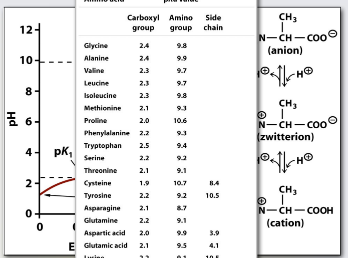
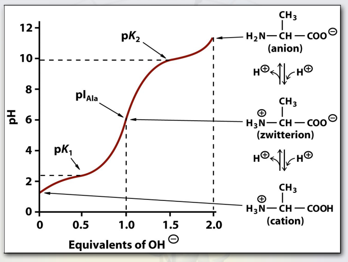


TABLE 3.2 pK_a values of acidic and basic constituents of free amino acids at 25°C

Amino acid	pKa value		
	Carboxyl group	Amino group	Side chain
Glycine	2.4	9.8	
Alanine	2.4	9.9	
Valine	2.3	9.7	
Leucine	2.3	9.7	
Isoleucine	2.3	9.8	
Methionine	2.1	9.3	
Proline	2.0	10.6	
Phenylalanine	2.2	9.3	
Tryptophan	2.5	9.4	
Serine	2.2	9.2	
Threonine	2.1	9.1	
Cysteine	1.9	10.7	8.4
Tyrosine	2.2	9.2	10.5
Asparagine	2.1	8.7	
Glutamine	2.2	9.1	
Aspartic acid	2.0	9.9	3.9
Glutamic acid	2.1	9.5	4.1
Lysine	2.2	9.1	10.5
Arginine	1.8	9.0	12.5
Histidine	1.8	9.3	6.0

26-2

The Amino Acids as Acids & Bases



26-3

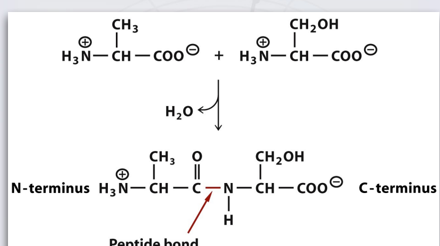
The Peptide Bond

Amino acids are joined together by peptide (amide) bonds.

- A peptide bond is formed from the condensation of an α-amino group from one amino acid with the α-carboxyl group of another amino acid.

27-1

The Peptide Bond



27-2

The Peptide Bond

Amino acids are joined together by peptide (amide) bonds.

- A peptide bond is formed from the condensation of an α-amino group from one amino acid with the α-carboxyl group of another amino acid.

Chem 352, Lecture 3 - Amino Acids and Protein Primary Structure 27

27-3

Peptides

Some small peptide have important biological activities.

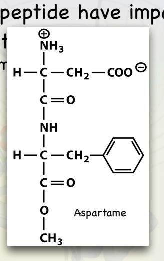
- peptide hormones
- toxins
- sweeteners

28-1

Peptides

Some small peptide have important biological act

- peptide horm
- toxins
- sweeteners



28-2

Peptides

Some small peptide have important biological activities.

- peptide hormones
- toxins
- sweeteners

28-3

Peptides

Problem:
Sketch the titration curve for the dipeptide cysteinylaspartic acid and draw the structures of the predominant species that exists at each of the endpoints. Confirm your answers using Marvinsketch.

29

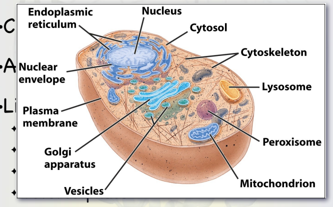
Isolation of Peptides and Proteins

- Cell disruption
- Centrifugation
- Ammonium sulfate precipitation
- Liquid chromatography
  - ion exchange
  - size exclusion (gel filtration)
  - affinity chromatography
  - reverse phase.

30-1

Isolation of Peptides and Proteins

•Cell disruption



30-2

Isolation of Peptides and Proteins

•Cell disruption

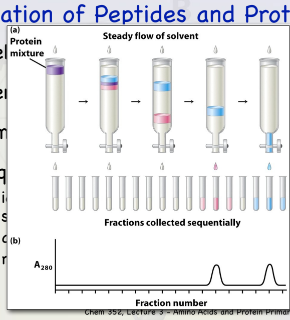
•Centrifugation

•Ammonium sulfate precipitation

•Liquid chromatography
- ion exchange
- size exclusion (gel filtration)
- affinity chromatography
- reverse phase.

30-3

Isolation of Peptides and Proteins



30-4

Isolation of Peptides and Proteins

•Cell disruption

•Centrifugation

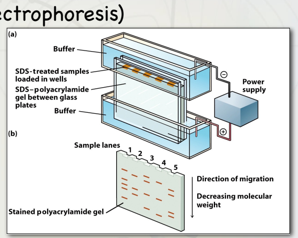
•Ammonium sulfate precipitation

•Liquid chromatography
- ion exchange
- size exclusion (gel filtration)
- affinity chromatography
- reverse phase.

30-5

Protein Isolation and Analysis

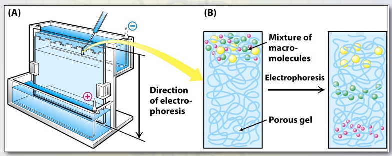
•SDS-PAGE (SDS-polyacrylamide gel electrophoresis)



31-1

Protein Isolation and Analysis

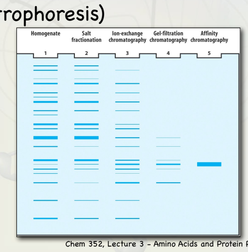
•SDS-PAGE (SDS-polyacrylamide gel electrophoresis)



31-2

Protein Isolation and Analysis

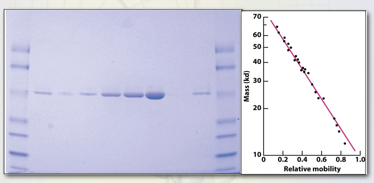
•SDS-PAGE (SDS-polyacrylamide gel electrophoresis)



31-3

Protein Isolation and Analysis

•SDS-PAGE (SDS-polyacrylamide gel electrophoresis)

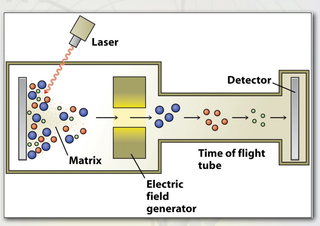


31-4

Protein Characterization

•Mass spectroscopy

MALDI Ionization



32-1

Protein Characterization

•Mass spectroscopy

ESI Ionization

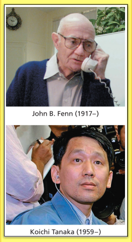
32-2

Protein Characterization

- Mass spectroscopy

ESI Ionization

MALDI Ionization



Nobel Prize in Chemistry, 2002

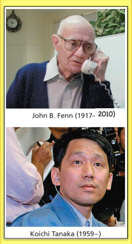
32-3

Protein Characterization

- Mass spectroscopy

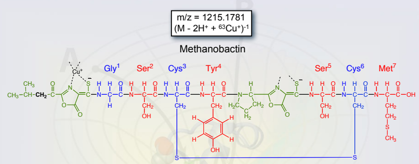
ESI Ionization

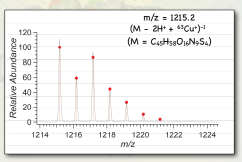
MALDI Ionization



Nobel Prize in Chemistry, 2002

32-4





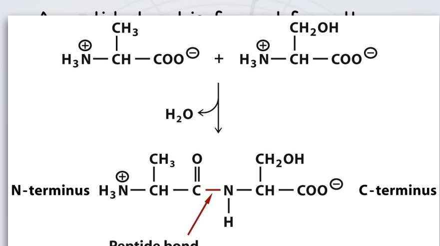
33

Protein Primary Structure

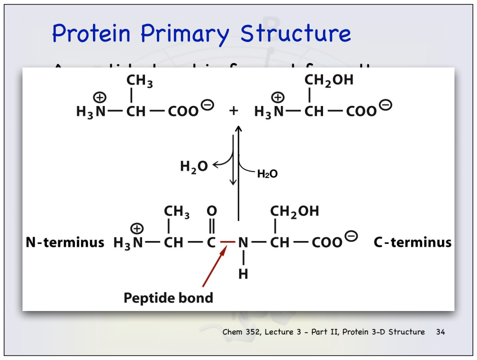
A peptide bond is formed from the condensation of an α-amino group from one amino acid with the α-carboxyl group of another amino acid.

34-1

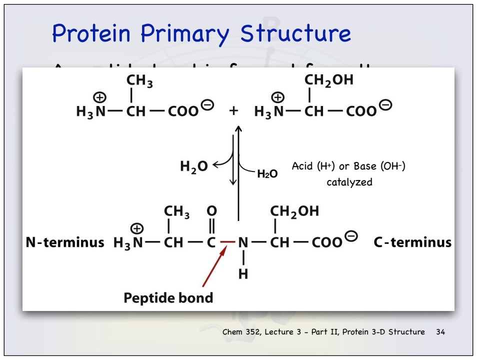
Protein Primary Structure



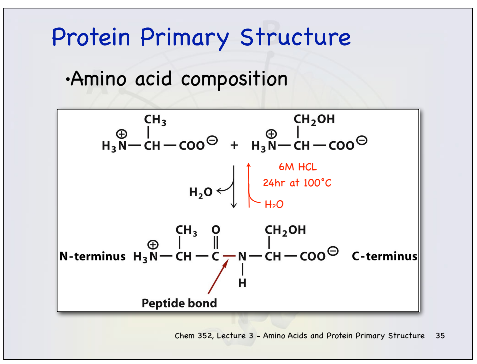
34-2



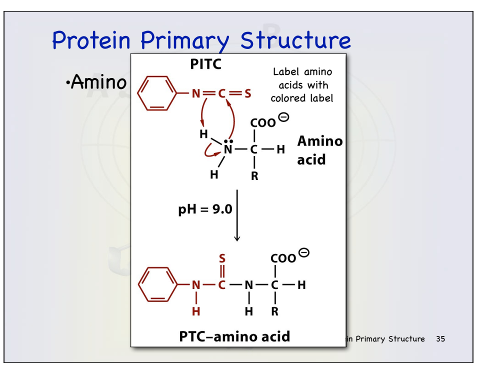
34-3



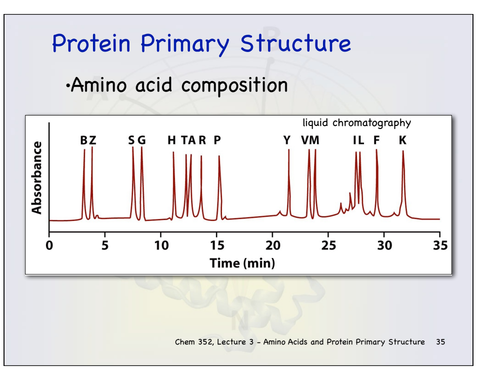
34-4



35-1



35-2



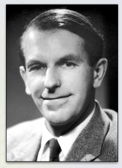
35-3

Protein Primary Structure

36

Fredrick Sanger was the first person to sequence a complete protein

- Insulin (1953)



1956 Nobel Prize in Chemistry

GIVEQCCASVCSLYQLENYCN — A chain (21 amino acids)

FVNQHLCGSHLVEALYLVCGERGFFYTPKA — B chain (30 amino acids)

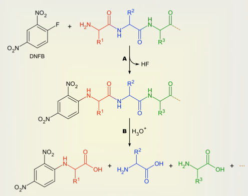
Figure 1.—The structure of bovine insulin.

Perspectives on Genetics: Antony Stretton, "The First Sequence: Fred Sanger and Insulin", Genetics 2002, 162, 527-532.

Protein Primary Structure

37

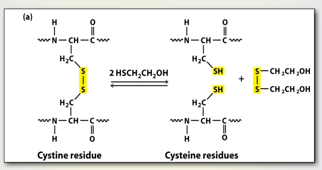
- Sanger's Reagent
  - (DNFB, 2,4-dinitrofluorobenzene)



Protein Primary Structure

38

- Cleavage of disulfide bonds with β-mercaptoenthanol.



Protein Primary Structure

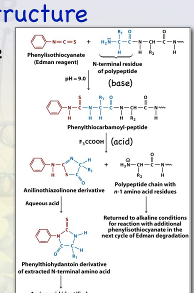
39-1

- Amino acid sequence
  - Edman Degradation

Protein Primary Structure

39-2

- Amino acid sequence
  - Edman Degradation



Protein Primary Structure

Amino acid sequence

- Edman Degradation
- Can only sequence up to around 30 amino acids at a time, therefore, for larger proteins, the polypeptide is cleaved into smaller segments.
  - **CNBr** (cyanogen bromide) cleaves at Met
  - **Trypsin** protease cleaves at Lys & Arg (+)
  - **Chymotrypsin** cleaves at Phe, Tyr & Trp (aromatic)

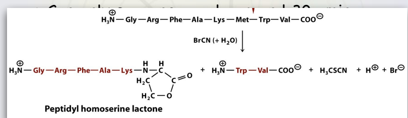
Chem 352, Lecture 3 - Amino Acids and Protein Primary Structure 40

40-1

Protein Primary Structure

Amino acid sequence

- Edman Degradation



  - **Trypsin** protease cleaves at Lys & Arg (+)
  - **Chymotrypsin** cleaves at Phe, Tyr & Trp (aromatic)

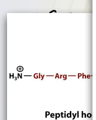
Chem 352, Lecture 3 - Amino Acids and Protein Primary Structure 40

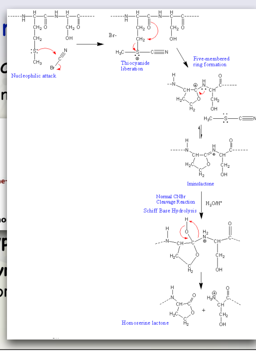
40-2

Protei

Amino a

- Edman





  - Tryp ... (+)
  - Chy ... rp (aro

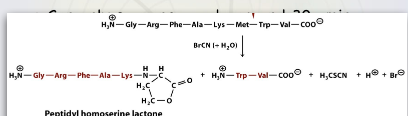
...ry Structure 40

40-3

Protein Primary Structure

Amino acid sequence

- Edman Degradation



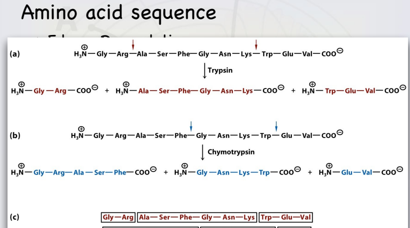
  - **Trypsin** protease cleaves at Lys & Arg (+)
  - **Chymotrypsin** cleaves at Phe, Tyr & Trp (aromatic)

Chem 352, Lecture 3 - Amino Acids and Protein Primary Structure 40

40-4

Protein Primary Structure

Amino acid sequence



Chem 352, Lecture 3 - Amino Acids and Protein Primary Structure 40

40-5

Protein Primary Structure

41

Question:

You have isolated a decapeptide called FP, which has anticancer activity. Determine the sequence of the peptide using the following information:

a. One cycle of Edman degradation of intact FP yields 2 mol of PTH-aspartate per mole of FP.
b. Treatment of a solution of FP with 2-mercaptoethanol followed by addition of trypsin yields three peptides with the composition (Ala, Cys,Phe), (Arg,Asp), and (Asp,Cys,Gly,Met,Phe). The intact (Ala,Cys, Phe) peptide yields PTH-cysteine in the first cycle of Edman degradation.
c. Treatment of 1 mol of FP with carboxypeptidase (which cleave the C-terminal residue from peptides) yields 2 mol of phenylalanine.
d. Treatment of the intact pentapeptide (Asp,Cys,Gly,Met,Phe) with CNBr yields two peptides with the composition (homoserine lactone, Asp) and (Cys,Gly,Phe). The (Cys,Gly,Phe) peptide yields PTH-glycine in the first cycle of Edman degradation.

Protein Primary Structure

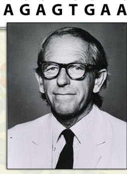
42-1

DNA ~~~ A A G A G T G A A C C T G T C ~~~

Protein Primary Structure

42-2

DNA ~~~ A A G A G T G A A C C T G T C ~~~



1956 Nobel Prize in Chemistry
1980 Nobel Prize in Chemistry

Protein Primary Structure

43-1

- Amino acid sequence
  - Reverse translating a DNA sequence

DNA ~~~ A A G A G T G A A C C T G T C ~~~

Protein ~~~ Lys — Ser — Glu — Pro — Val ~~~

Protein Primary Structure

43-2

- Amino acid sequence
  - Reverse translating a DNA sequence

DNA ~~~ ... T G T C ~~~

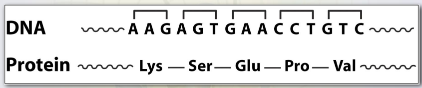
Protein ~~~ ... — Val ~~~

	U	C	A	G	
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	STOP	STOP	A
	Leu	Ser	STOP	Trp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

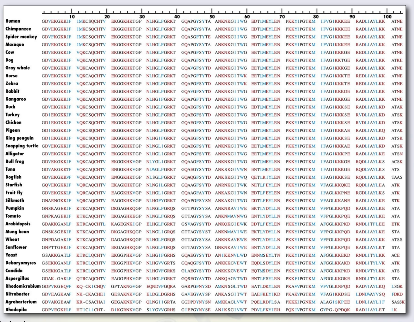
43-3

Protein Primary Structure

- Amino acid sequence
  - Reverse translating a DNA sequence

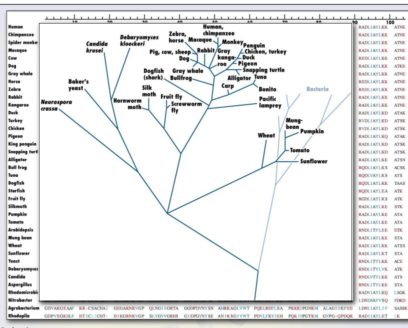


44-1



Cytochrome c sequence

44-2



Cytochrome c sequence

45-1

Protein Primary Structure

Question:

The functional diversity of proteins results from the large number of possible polypeptides that can be built using the 20 different amino acids

Question: What is the minimum mass it would take to construct one molecule each of all of the possible polypeptides that contain 100 amino acids?

45-2

Protein Primary Structure

Question:

The functional diversity of proteins results from the large number of possible polypeptides that can be built using the 20 different amino acids

Question: What is the minimum mass it would take to construct one molecule each of all of the possible polypeptides that contain 100 amino acids?

45-3

Protein Primary Structure

Number of polypeptides (20^{100})	1.26×10^{130}
Avg. Mass of each polypeptide	1.83×10^{-22} g
Total mass needed	2.32×10^{108} g
Number of Earths	3.9×10^{80}
Number of Suns	1.2×10^{75}
Number of Galaxies	9.7×10^{29}

45-4

Protein Primary Structure

Question:

The functional diversity of proteins results from the large number of possible polypeptides that can be built using the 20 different amino acids

Question: What is the minimum mass it would take to construct one molecule each of all of the possible polypeptides that contain 100 amino acids?

46

Next up

Lecture 3, Part II - Protein Structure and Function

- Read Chapter 4 of Moran et al.