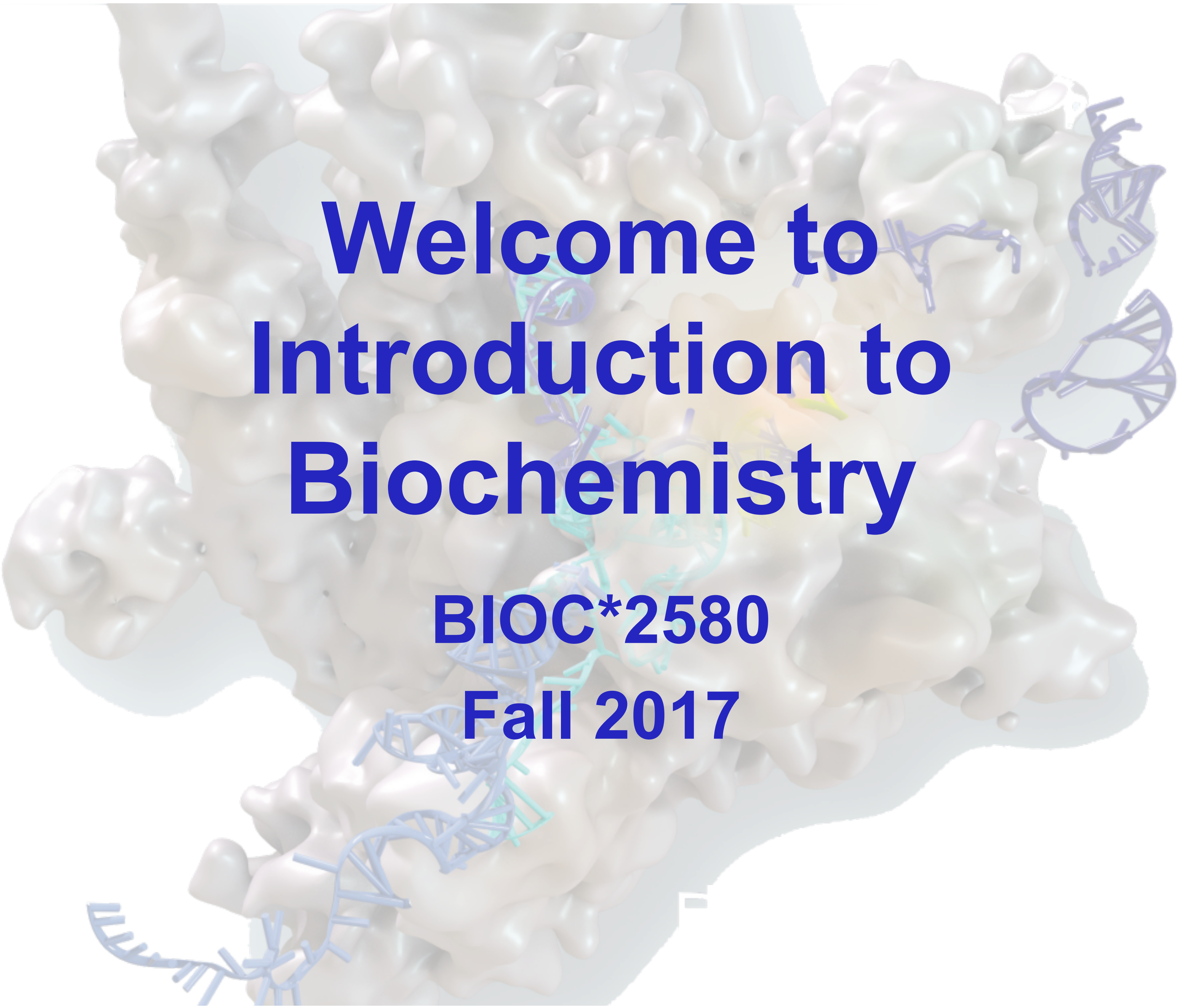


Welcome to Introduction to Biochemistry

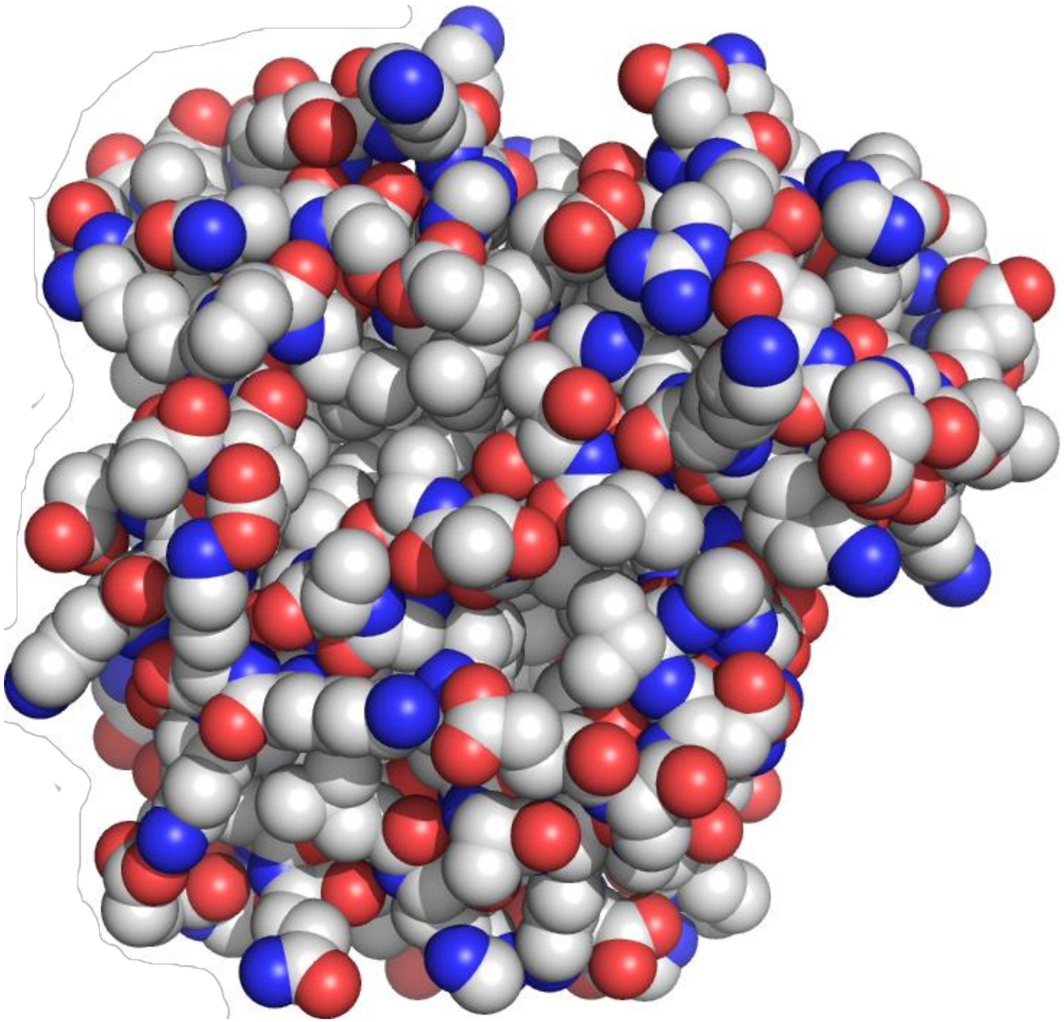
BIOC*2580

Fall 2017

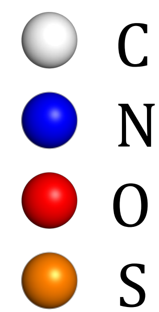
Molecules we study in biochemistry

- Small molecules
  - sugars, amino acids, nucleotides, carboxylic acid derivatives
  - act as building blocks for macromolecules
- Macromolecules
  - Proteins – chains of amino acids
  - Polysaccharides – chains of simple sugars
  - Nucleic acids – chains of nucleotides

A macromolecule



myoglobin is a **protein** that stores O_2 in muscle tissue



How large is a protein molecule?

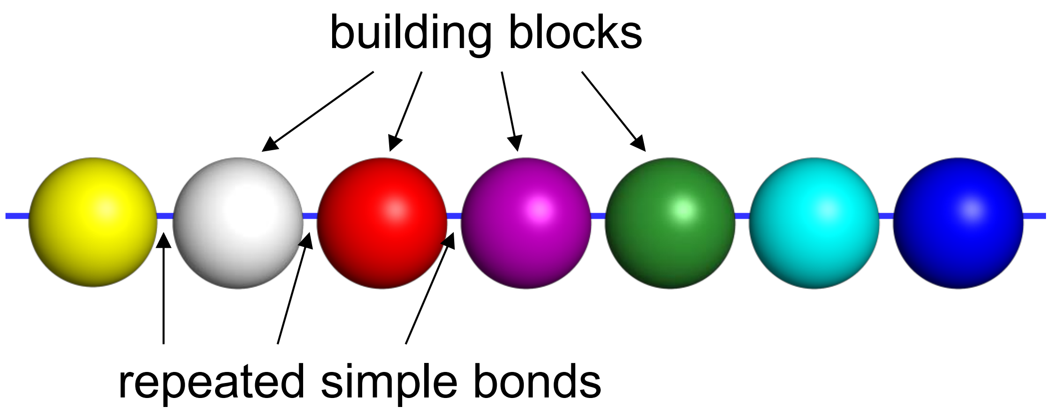
- Most proteins: **10,000 to 100,000 g mol^{-1}**
- Protein size is expressed in <u>kiloDaltons (kDa)</u>

  1 Dalton (Da) = 1 g mol^{-1} (mass of H atom)

  1 kDa = 1000 g mol^{-1}

- Myoglobin is 16.5 kDa – small protein
- P-glycoprotein is 170 kDa – large protein

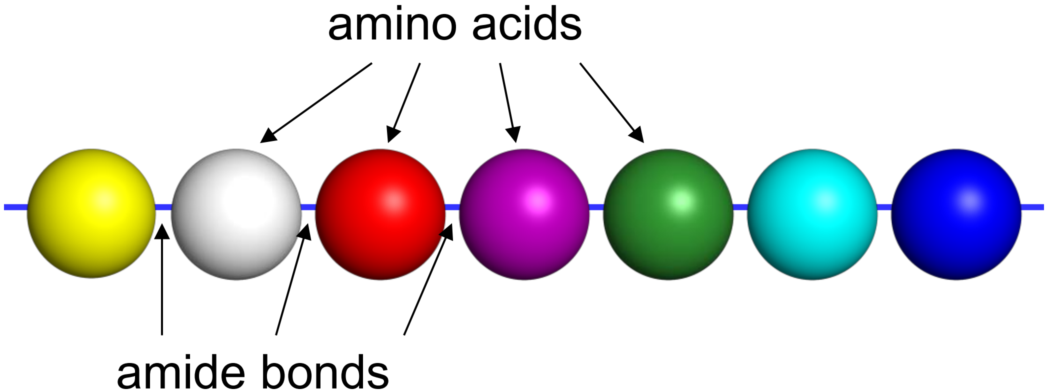
The building block principle of macromolecular structure



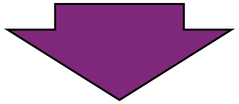
Understand the building blocks

Understand the macromolecules

Proteins are made of amino acids



Understand amino acids



Understand proteins

Proteins are made of amino acids

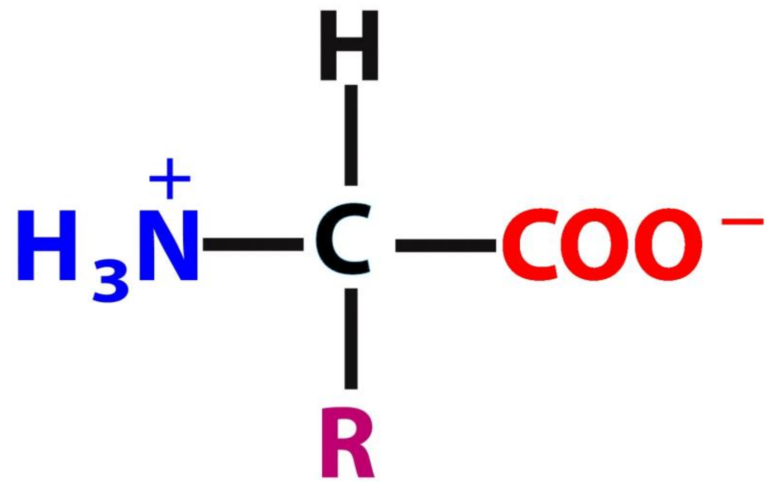
- linear chains of **amino acids**
- linked by **peptide bonds** (type of amide bond)

- Each protein has:
  - **unique sequence of different amino acids**
  - a well defined **size and structure**

- Proteins have diverse functions including:
  - **catalyzing reactions (enzymes)**
  - forming complex subcellular structures

Basic amino acid structure

- Each amino acid has an <u>amino</u> **group** and a <u>carboxylate</u> **group**
- Each amino acid has a different <u>side chain R</u>
- 20 different amino acids are found in proteins (see Lehninger p79 (75))



Peptide bonds

- Condensation involves removal of **H_2O** from the units being linked
- Hydrolysis regenerates the original carboxylic acid and amino group
- The **C=O group** of the amide is the **point of weakness** allowing H_2O attack

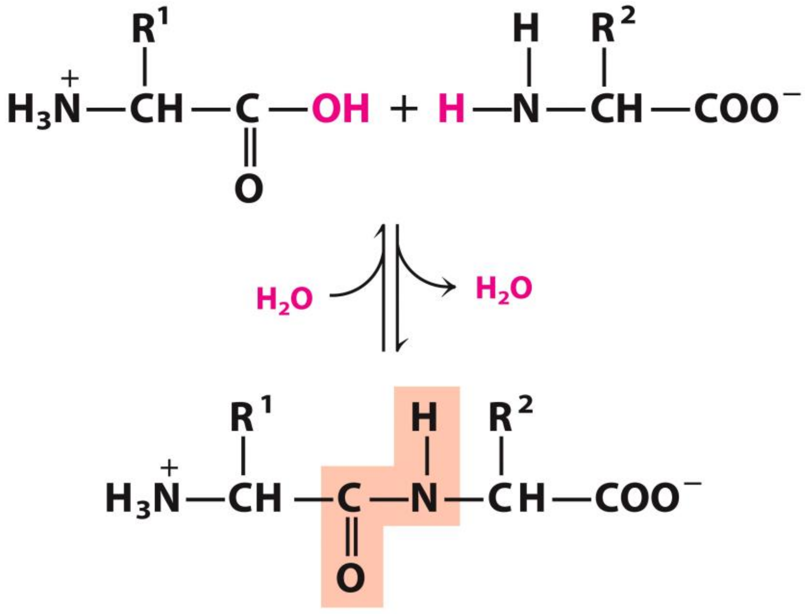
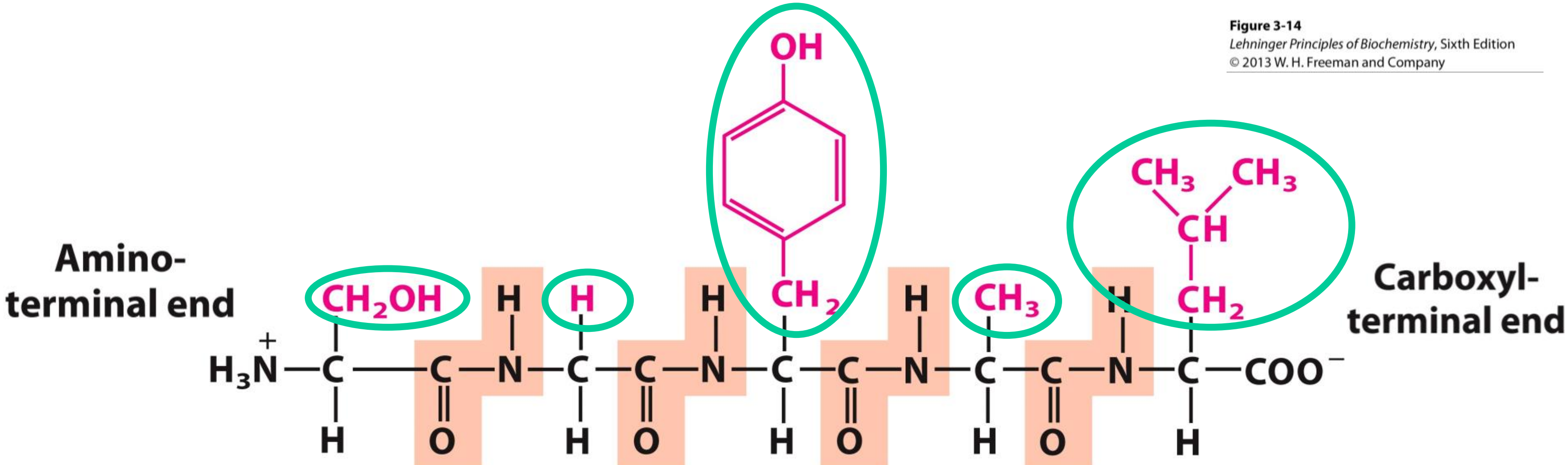


Figure 3-13
Lehninger Principles of Biochemistry, Sixth Edition

Large numbers of amino acids can be linked together to form a peptide chain



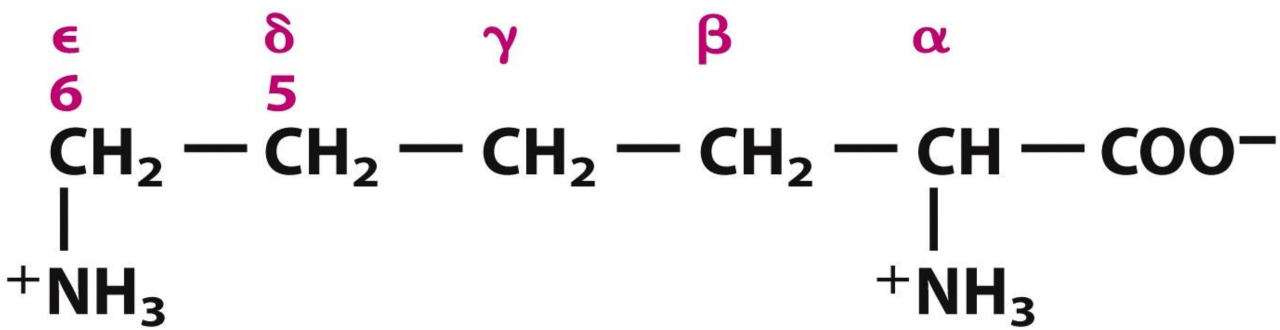
- The combination of different side chains R1, R2, R3, etc gives each protein its unique properties
- there are 153 amino acids in myoglobin (16.5 kDa)

Polypeptide – a chain with *many* amino acids, usually a complete protein
Greek *poly* = many

Oligopeptide – a chain with a *few* amino acids, usually a fragment
Greek *oligo* = a few

Amino acid side chain structure

- Carbon atoms of the amino acid core are identified by Greek letters
- The ***α***-carbon is the central backbone atom
- The ***β***-carbon is the first atom of the side chain, the **γ-carbon** is the second, etc
- Functional groups may be linked to different core atoms:
  - **α-amino**
  - **ε-amino**



Lysine

The 20 natural amino acids

- Amino acids share a common backbone, but differ in the **side chain**
- You need to be able to **reproduce the structures of the amino acids** and know their 1- and 3-letter codes
- You need to be able to associate particular properties with each amino acid
  - Polarity
  - Charge
  - Hydrogen bonding ability
- Amino acids can be grouped according to **structures** or by **similar properties**

Alanine
Ala, A

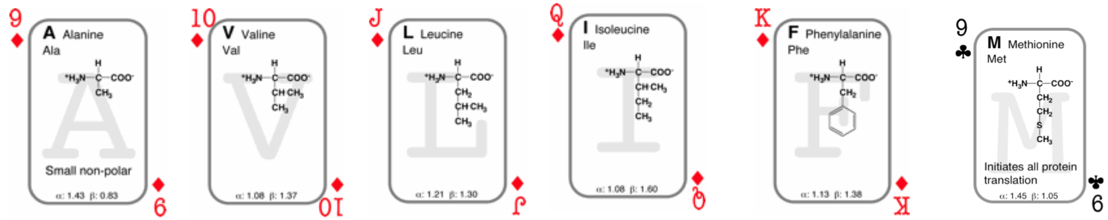
Valine
Val, V

Leucine
Leu, L

Isoleucine
Ile, I

Phenylalanine
Phe, F

Methionine
Met, M



6 with **very non-polar** side chains

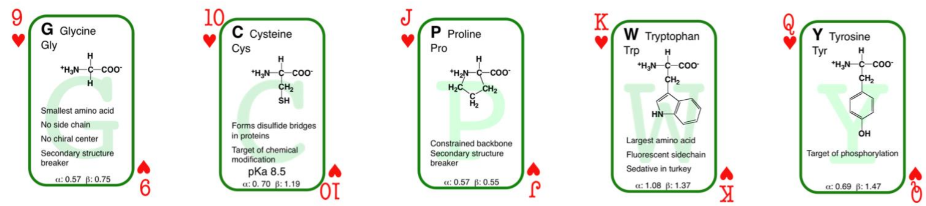
Glycine
Gly, G

Cysteine
Cys, C

Proline
Pro, P

Tryptophan
Trp, W

Tyrosine
Tyr, Y



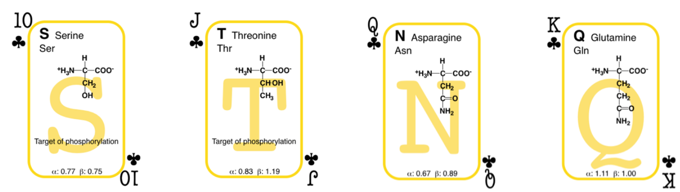
5 with **moderately non-polar** side chains

Serine
Ser, S

Threonine
Thr, T

Asparagine
Asn, N

Glutamine
Gln, Q

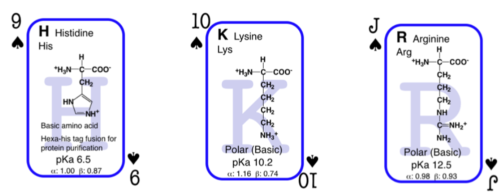


4 with **polar** but **uncharged** side chains

Histidine
His, H

Lysine
Lys, K

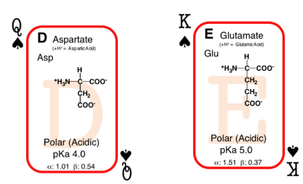
Arginine
Arg, R



3 with **positively charged** side chains (very polar)

Aspartate
Asp, D

Glutamate
Glu, E

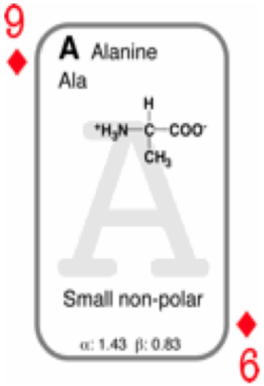


2 with **negatively charged** side chains (very polar)

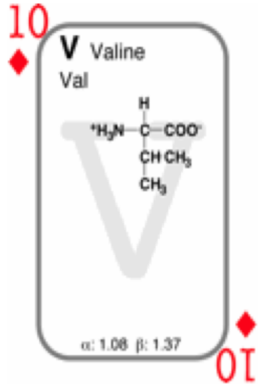
"inverted pyramid"
6 5 4 3 2

6 with **very non-polar** side chains

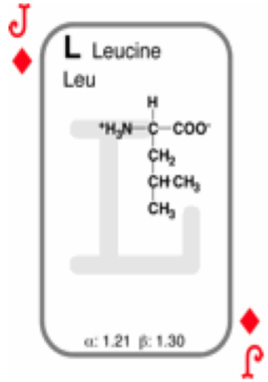
Alanine
Ala, A



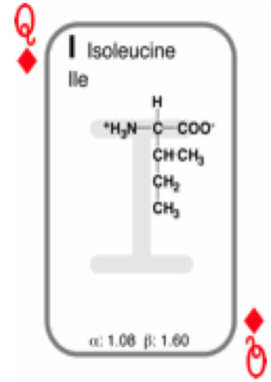
Valine
Val, V



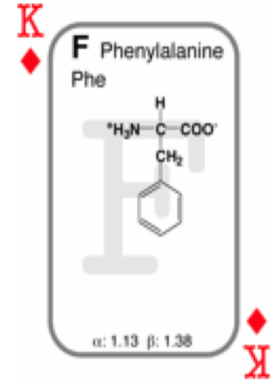
Leucine
Leu, L



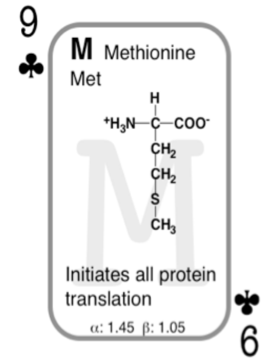
Isoleucine
Ile, I



Phenylalanine
Phe, F

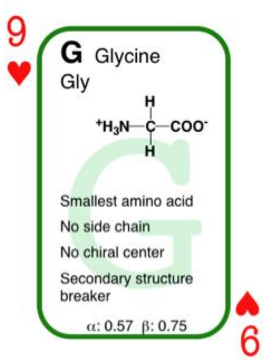


Methionine
Met, M

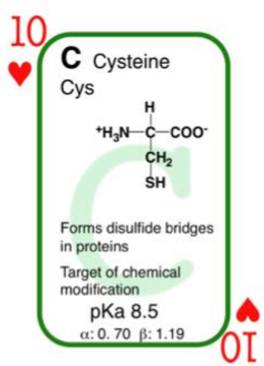


5 with **moderately non-polar** side chains

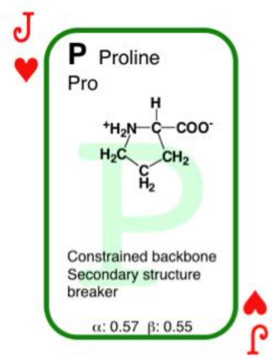
Glycine
Gly, G



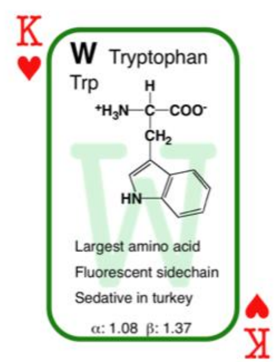
Cysteine
Cys, C



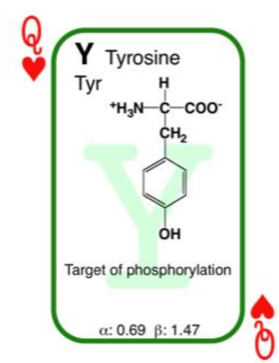
Proline
Pro, P



Tryptophan
Trp, W

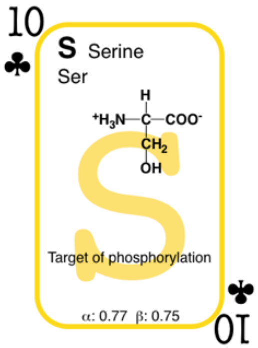


Tyrosine
Tyr, Y

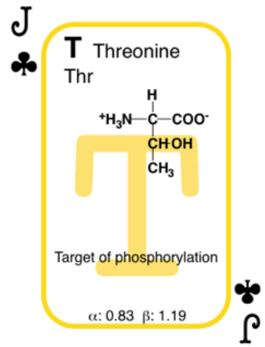


4 with **polar** but **uncharged** side chains

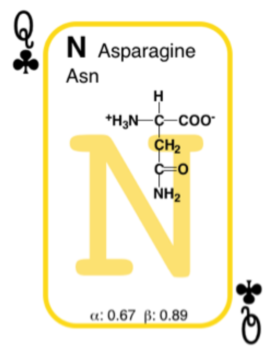
Serine
Ser, S



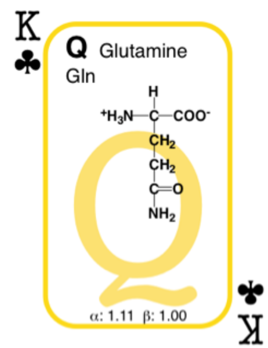
Threonine
Thr, T



Asparagine
Asn, N

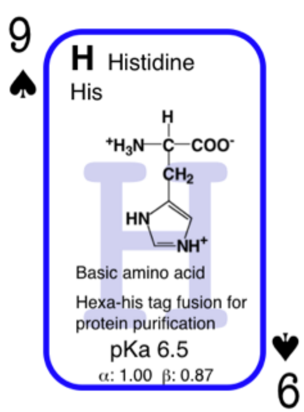


Glutamine
Gln, Q

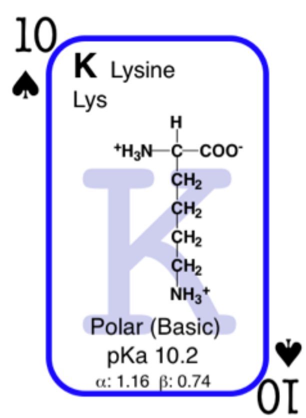


3 with **positively charged** side chains **(very polar)**

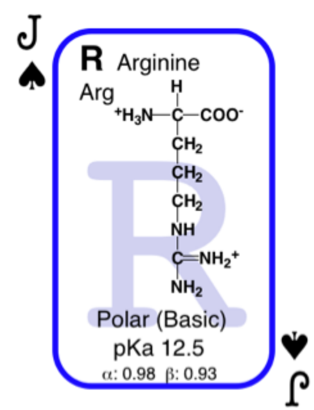
Histidine
His, H



Lysine
Lys, K



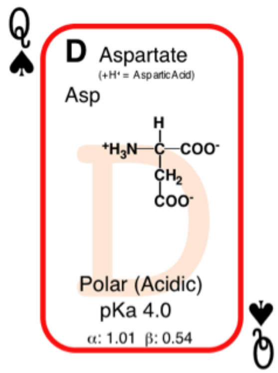
Arginine
Arg, R



2 with **negatively charged** side chains **(very polar)**

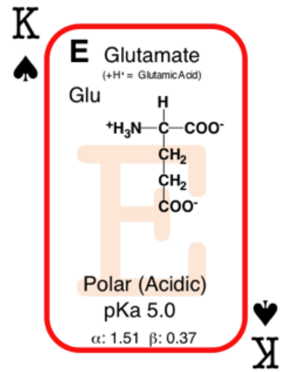
Aspartate

Asp, D



Glutamate

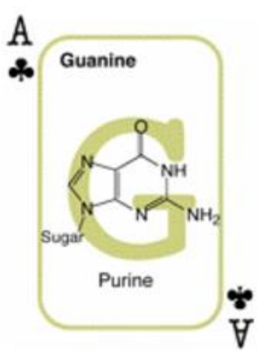
Glu, E



4 nucleotides in DNA (second half of the course)

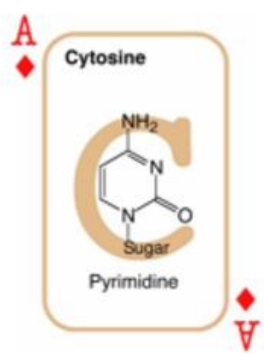
Guanine

G



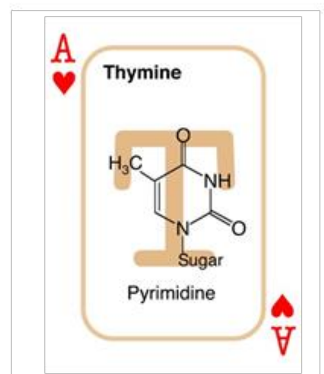
Cytosine

C



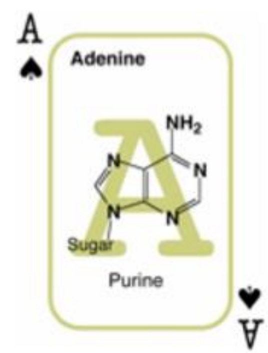
Thymine

T



Adenine

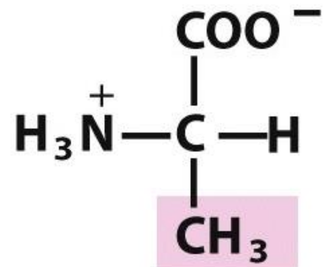
A



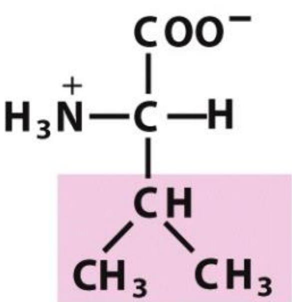
Do not need to memorize the structures of the nucleotide bases for the midterm. They will be tested in the final exam.

Amino acids with very non-polar side chains
Ala, Val, Leu, Ile, Met, Phe

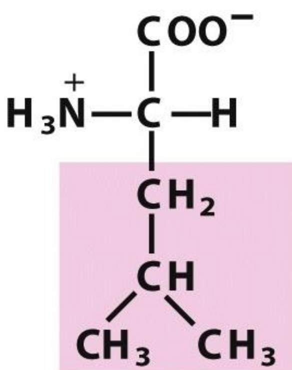
- The side chains are dominated by **hydrocarbon**, and consist only of **C-C** and **C-H** bonds
- Hydrocarbon is __non-polar__ and __hydrophobic__ (or water avoiding)



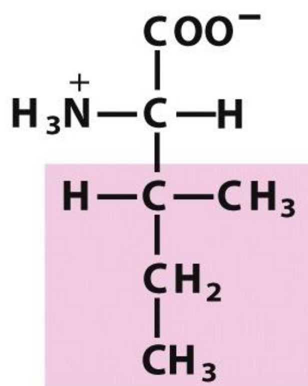
Ala (A)



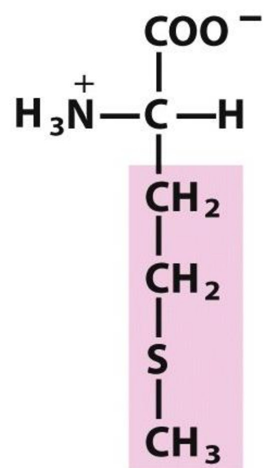
Val (V)



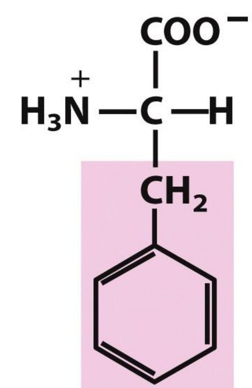
Leu (L)



Ile (I)



Met (M)



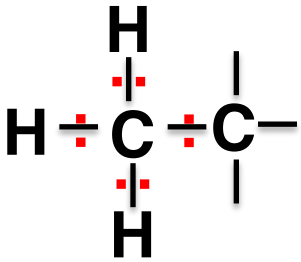
Phe (F)
(*F*enylalanine)

Polar and non-polar properties

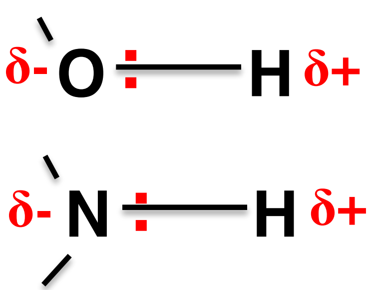
- Polarity is a consequence of atoms having different **electronegativity** or tendency to hold bonding electrons

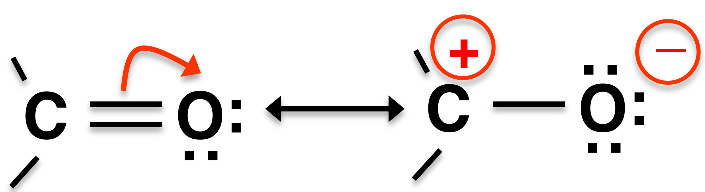
O > N > S > C ≅ H

- Atoms with similar electronegativity share bonding electrons equally, e.g. C-C, C-H, and are **non-polar**



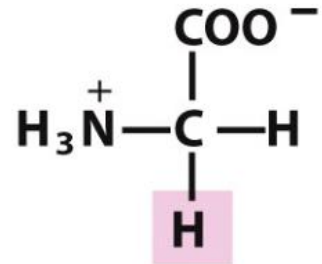
- Pairs of atoms with different electronegativity distribute bonding electrons unequally – more electronegative atoms such as **O** or **N** get greater than 50% share, and this leads to unbalanced charges and **polar** bonds



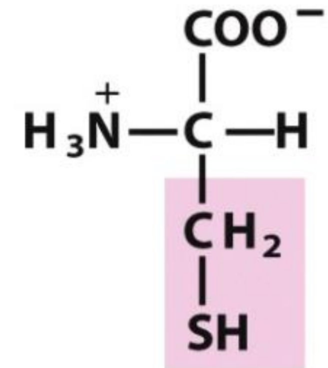


Moderately non-polar: Gly, Cys, Pro, Tyr, Trp

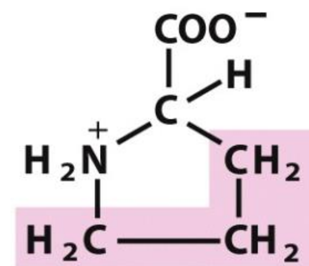
- **Glycine** has single H atom as side chain, not enough to be very non-polar
  - Hydrophobicity is related to the number of CH, CH_2 or CH_3 groups present
- **Cysteine** contains the slightly polar SH group
- **Proline** is unique because the side chain links to α-N as well as to α-C. The polar N moderates the non-polar hydrocarbon.
- **Tyrosine** has a single polar group that partly offsets the very non-polar benzene ring. **Tryptophan** behaves similarly.



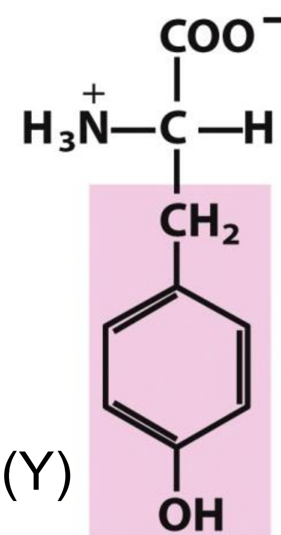
Gly (G)



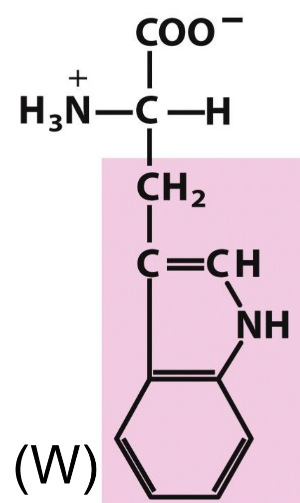
Cys (C)



Pro (P)



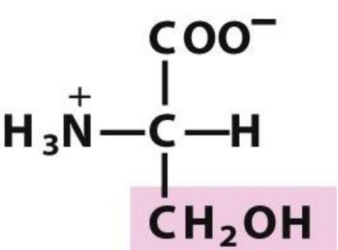
Tyr (Y)



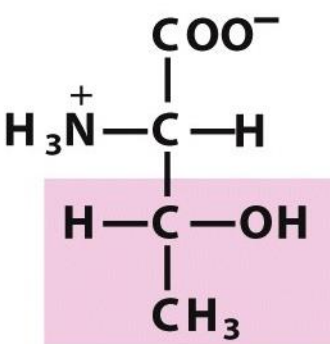
Trp (W)
(t*W*iptophan)

Amino acids with polar uncharged side chains
Ser, Thr, Asn, Gln

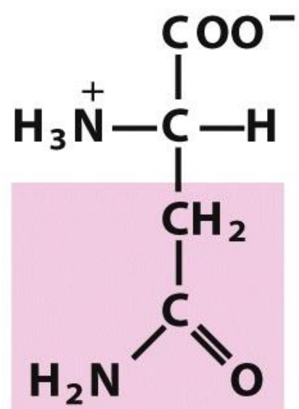
- **Serine** and **threonine** have side chains that include the polar hydroxyl group **-OH** (simple alcohol)
- **Asn** and **Gln** both contain the polar **amide** group
- These side chain groups do not gain or lose H^+ in aqueous solution at pH 7, so they are **uncharged**
- All four side chains act as **good** hydrogen bond **donors or acceptors**



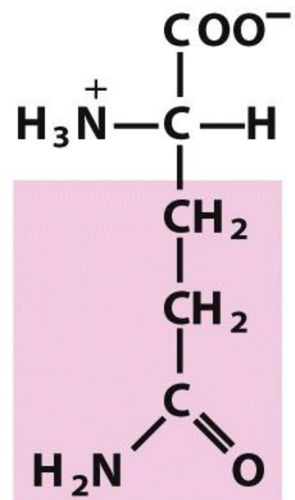
Ser (S)



Thr (T)



Asn (N)
(Asparagi*N*e)



Gln (Q)
(*Q*lutamine)

- **Hydrogen bonds** are electrostatic attractions between a H-bond **donor** and an **acceptor**

Highly polar –**OH** or –**NH** groups are good **H-bond donors**

R_1O-H - - :OR_2

δ– δ+

R_1O-H - - :NR_2

R_1N-H - - :NR_2

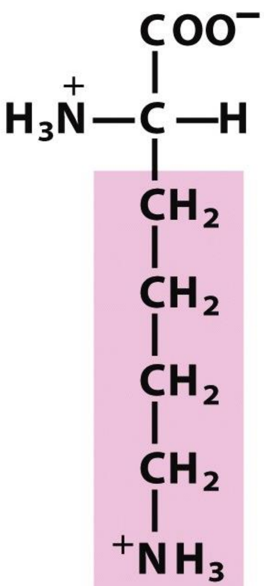
R_1N-H - - :OR_2

An **acceptor** is an electronegative atom with an **available** lone pair of electrons, such as O or N

The hydrogen bond (- -) is about **5-10%** as strong as a covalent bond, enough to make **molecule R_1 stick loosely to R_2** but not to form a permanent link

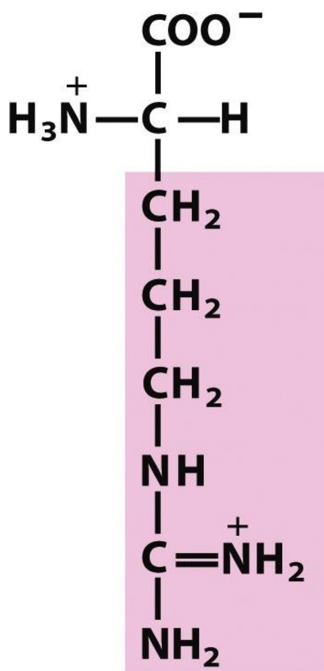
Positively charged side chains **His, Lys** and **Arg**

- These side chains contain weak bases that **gain H⁺** (become _protonated_______) and so are **positively** charged in aqueous solution at neutral pH
- Charge makes them _very polar_______, overriding the non-polar hydrocarbon chain



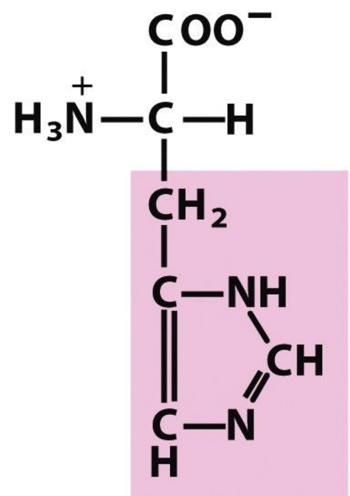
Lys (K)

(***K***-The letter before L)



Arg (R)

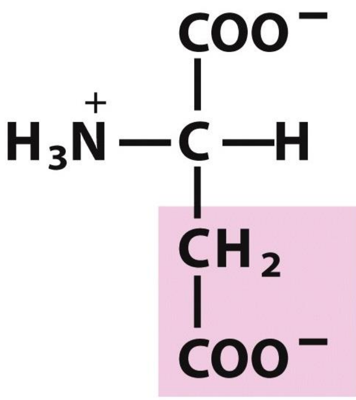
(a***R***ginine)



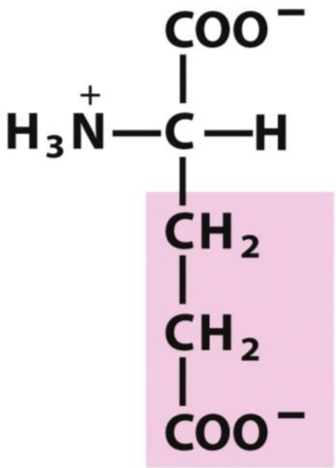
His (H)

Negatively charged side chains **Asp and Glu**

- Side chains have carboxylic acid groups R-COOH that lose H^+ (become **deprotonated**) at neutral pH
  - When deprotonated these are described as **carboxylate** groups R-**COO^-**
  - Carboxylate groups are **negative** and also **very polar**
- Asp side chain: $-CH_2-$**COO^-**
- Glu side chain: $-CH_2-CH_2-$**COO^-**



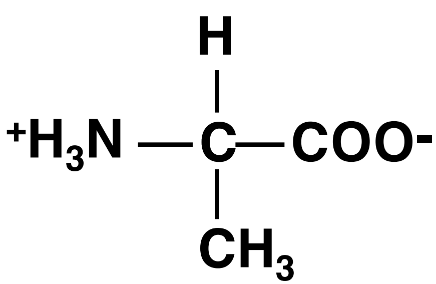
Asp (D)
(asparDate)

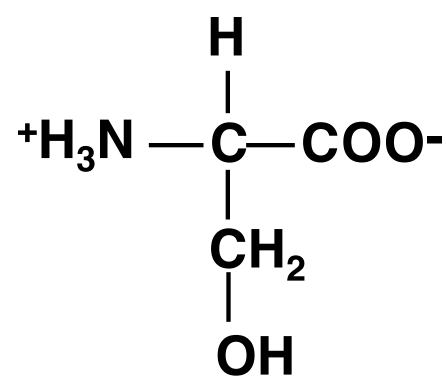


Glu(E)
glutamatE

Home work questions

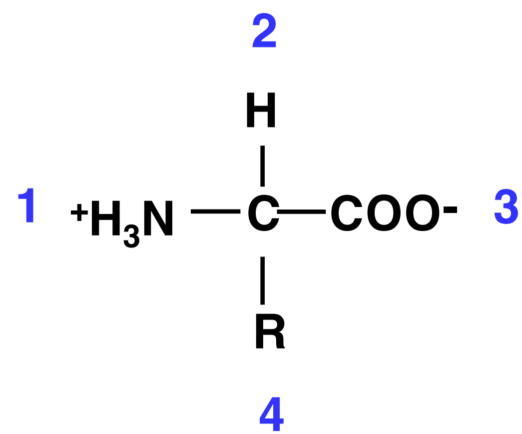
1. Shown below are the amino acids Ala and Ser. Join them together through a peptide bond to form the dipeptide alanylserine (Ala-Ser).





2. When an amino acid is described as being “charged”, which part of it are we referring to?

A. 1
B. 2
C. 3
D. 4
E. 1+3

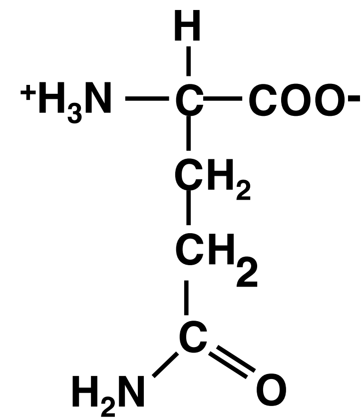


3. Which of the following amino acids can accept or donate hydrogen bonds on its side chain?

A. Val
B. Ala
C. Tyr
D. Phe

4. The amino acid is shown below is:

A. Very non-polar
B. Moderately non-polar
C. Polar and uncharged
D. Polar and +vely charged
E. Polar and –vely charged



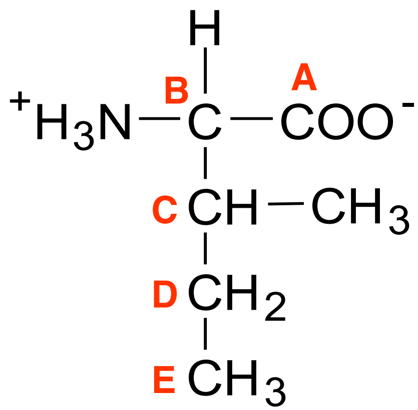
5. Which amino acid is present at the C-terminus of the following peptide?

QYTKLHSGAN

A. Asn
B. Ala
C. Gln
D. Glu

6. Which atom in the following diagram is the γ-carbon of the amino acid?

A.
B.
C.
D.
E.



Answers
1. -
2. D
3. C
4. C
5. A
6. D

Free amino acids are <u>weak electrolytes</u> due to their amino and carboxylate groups

- Normal biochemical processes occur close to pH 7 (physiological pH is 7.0-7.4)
- Groups such as <u>carboxylate</u> and <u>amino</u> groups gain or lose H^+ depending on availability of H^+ in solution
- pH expresses availability of H^+: **$pH = -\log_{10}[H^+]$**
  at pH = 7, $[H^+] = 10^{-7}$ M
- **The <u>Henderson-Hasselbalch</u> equation relates pH, pKa and the state of ionization of a given group**

$$pH = pK_a + \log \frac{[\text{deprotonated}]}{[\text{protonated}]}$$

The ionization state of the α-carboxylate group at pH 7

Typical $pK_a = 2.4 \pm 0.5$ for amino acid **α-carboxylate**

$$-COOH \rightleftharpoons -COO^- + H^+$$

$$pH - pK_a = \log \frac{[A^-]}{[HA]}$$

$$10^{(pH - pK_a)} = \frac{[COO^-]}{[COOH]}$$

At pH = 7.0 $$\frac{[COO^-]}{[COOH]} = 10^{(7.0-2.4)} = 40\,000$$

- The vast majority of molecules exist at pH 7 as carboxylate —COO-, NOT as carboxylic acid —COOH

The ionization state of the α-amino group at pH 7

Typical pK_a = 9.6 for amino acid α-amino groups

$$-NH_3^+ \rightleftharpoons -NH_2 + H^+$$

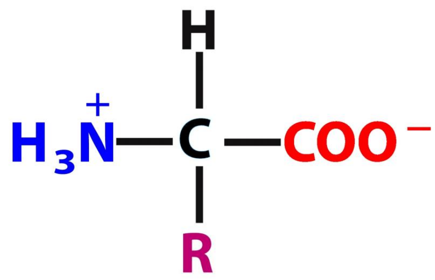
$$\frac{[\text{-NH}_2]}{[-\text{NH}_3^+]} = 10^{(\text{pH}-\text{pK}_a)}$$

At pH = 7.0, $$\frac{[\text{-NH}_2]}{[-\text{NH}_3^+]} = 10^{(7.0\text{-}9.6)} = 10^{-2.6} = 0.0025$$

- The vast majority of molecules exist as **protonated** __RNH3-__ rather than as __RNH3__

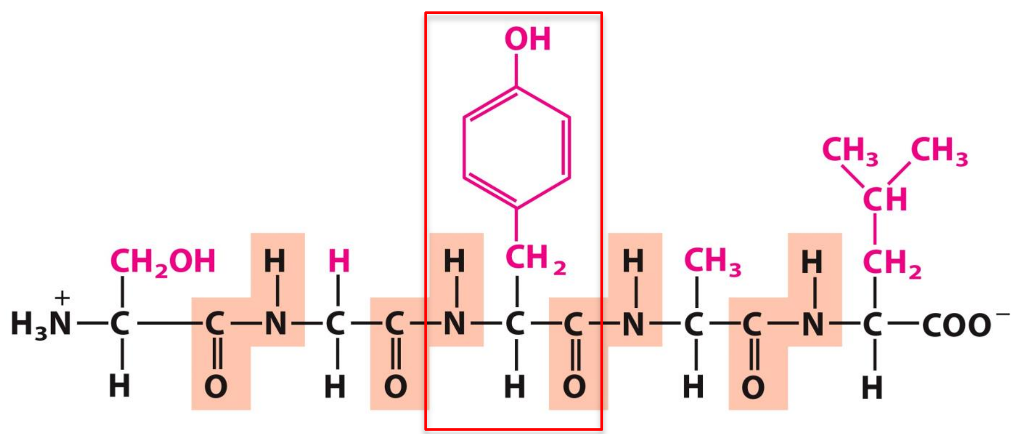
Charged state of amino acids at neutral pH

- The correct structure to represent an **individual amino** acid at neutral pH is



- But when the amino acid is **part of a peptide chain**, the α-amino groups and α-carboxylate groups are linked as uncharged amide bonds

x-NH-CHR-CO-x

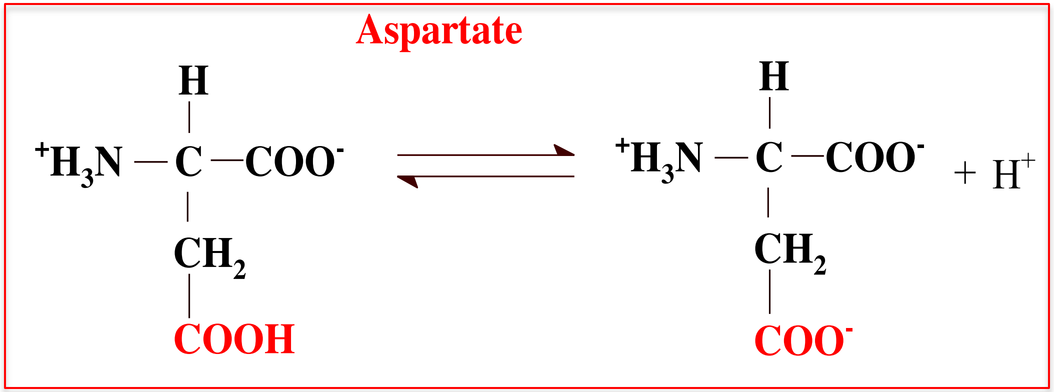


In addition to the α-amino and α-carboxylate groups, some R groups are also able to ionize

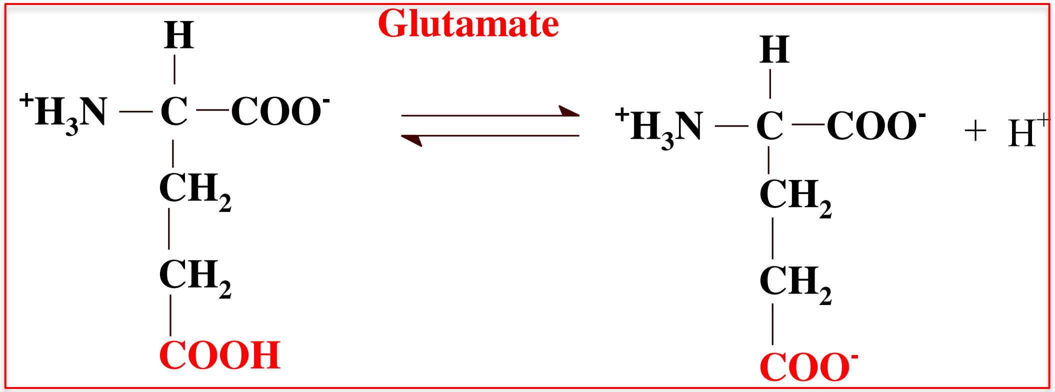
There are **seven** amino acids with **ionizable side chains**

1. Aspartate and glutamate contain carboxylate groups in their side chains (negatively charged at pH 7)

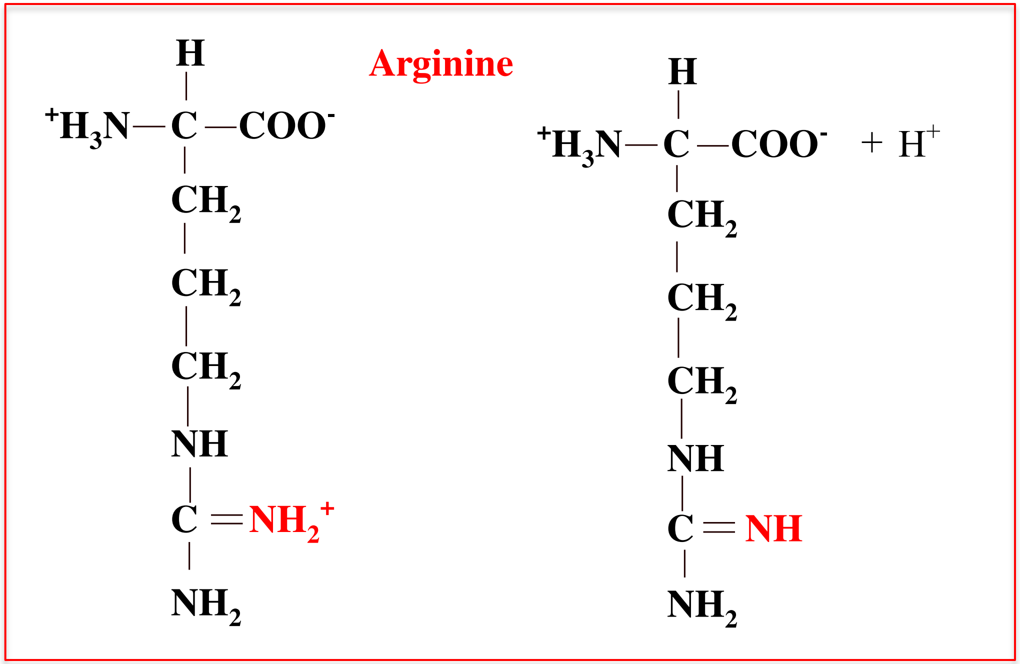
$$-COOH \rightleftharpoons -COO^- + H^+$$

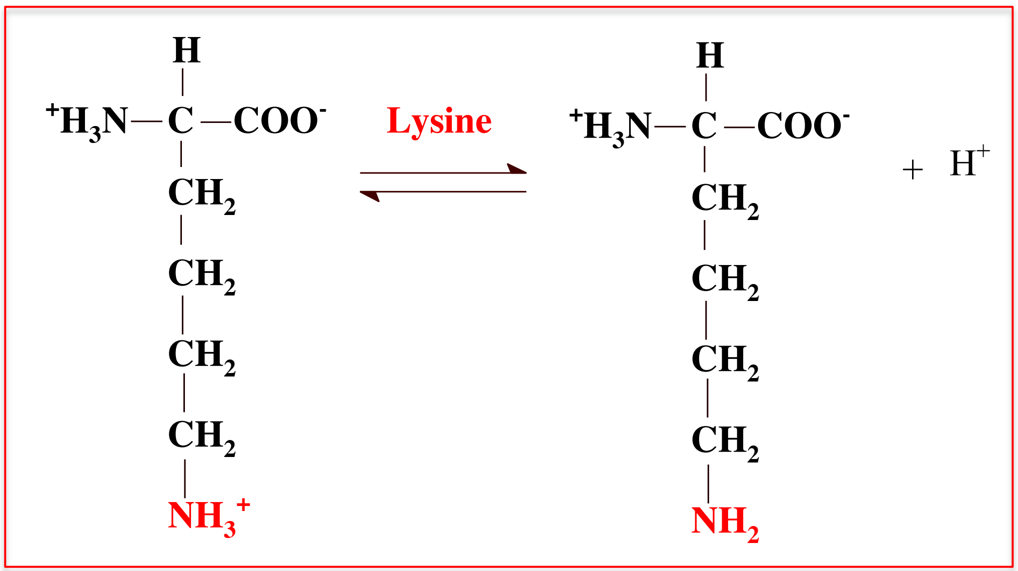


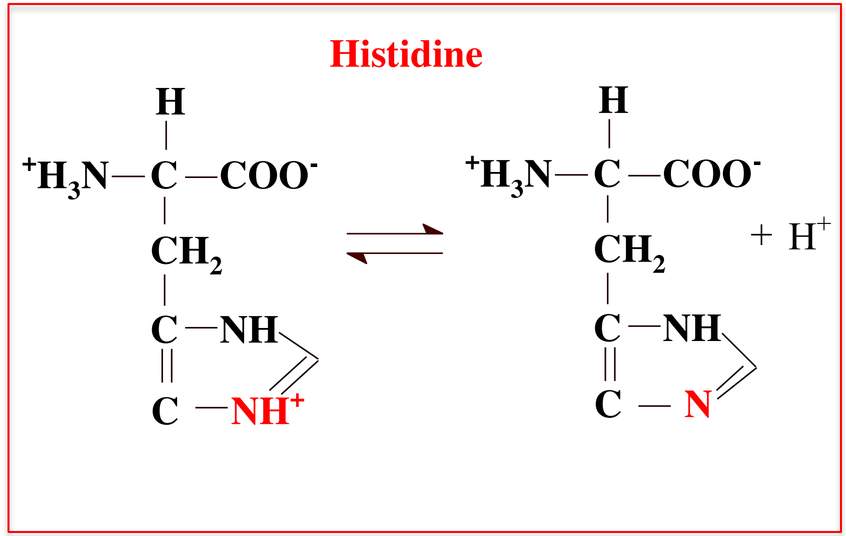
Side chain neutral (COOH) when protonated and -1 (COO-) when deprotonated



2. Arginine, Lysine and Histidine ionize on their side chain N atoms (positively charged at pH 7)

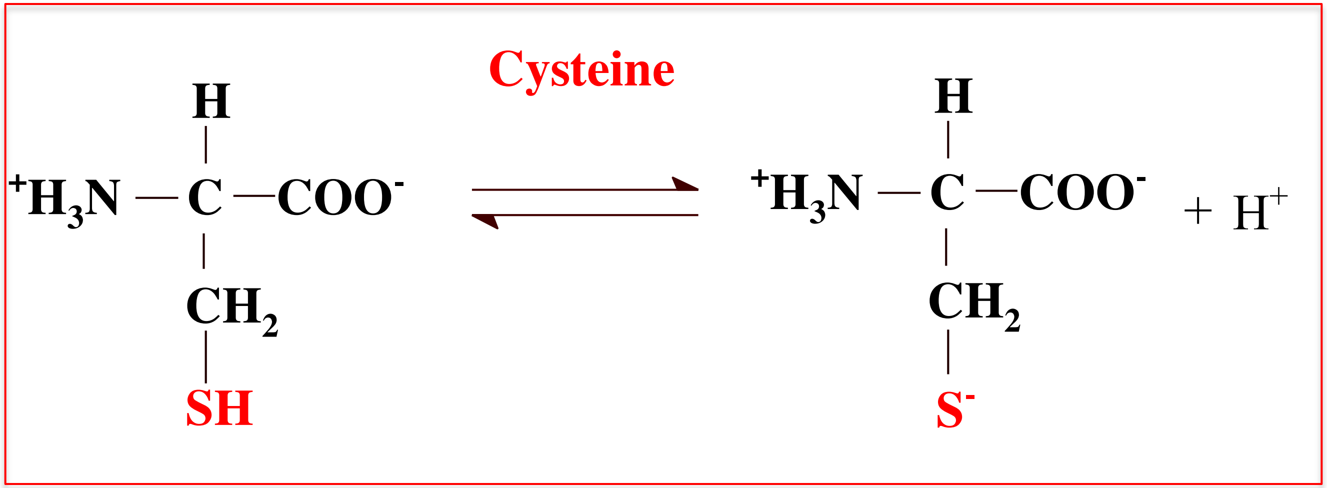


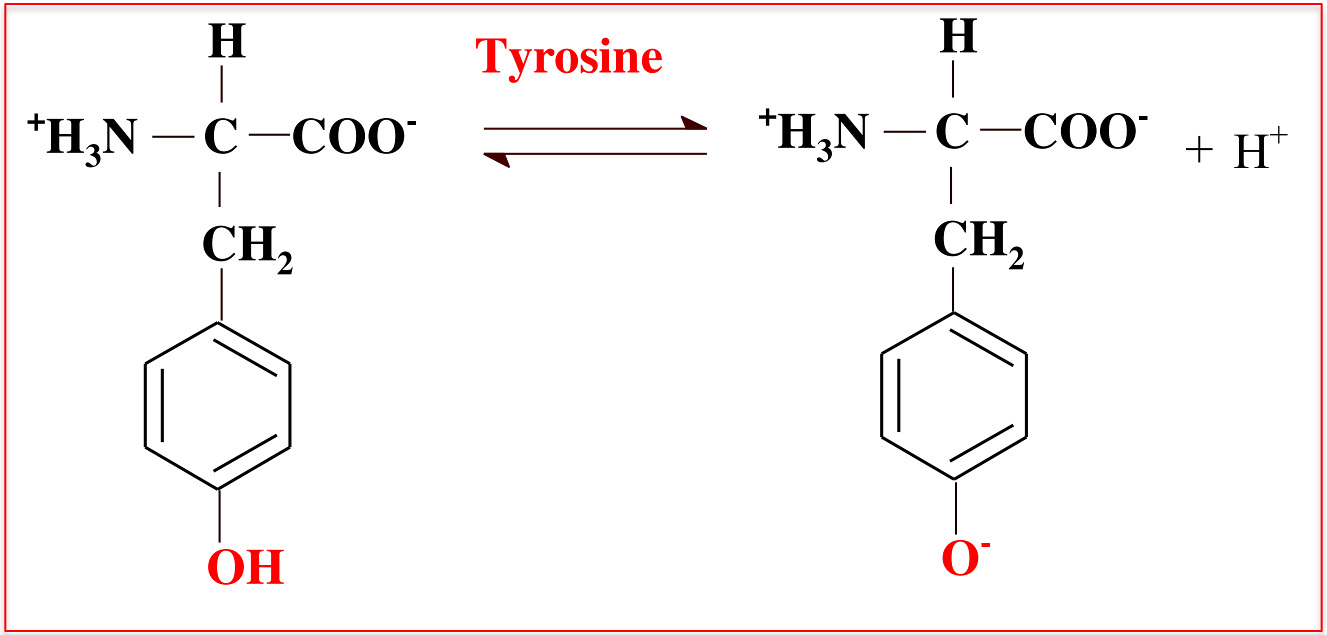




Side chains are +1 ($=NH_2^+$, $-NH_3^+$, $-NH^+$) when protonated and neutral ($=NH$, $-NH_2$, $-N$) when deprotonated

3. Cysteine and Tyrosine also have ionizable side chains (Neutral at pH 7)

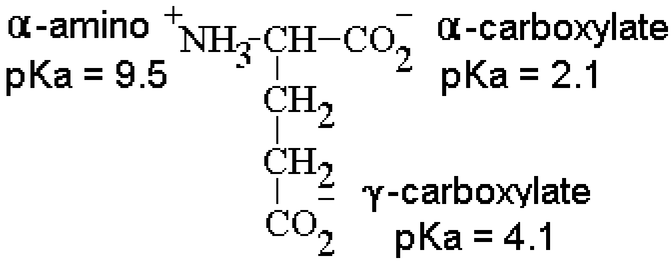


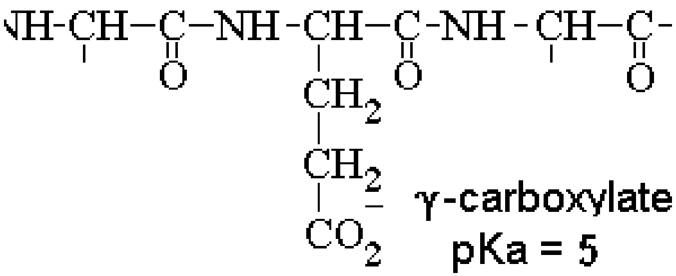


Side chain neutral (SH, OH) when protonated and -1 (S^-, O^-) when deprotonated

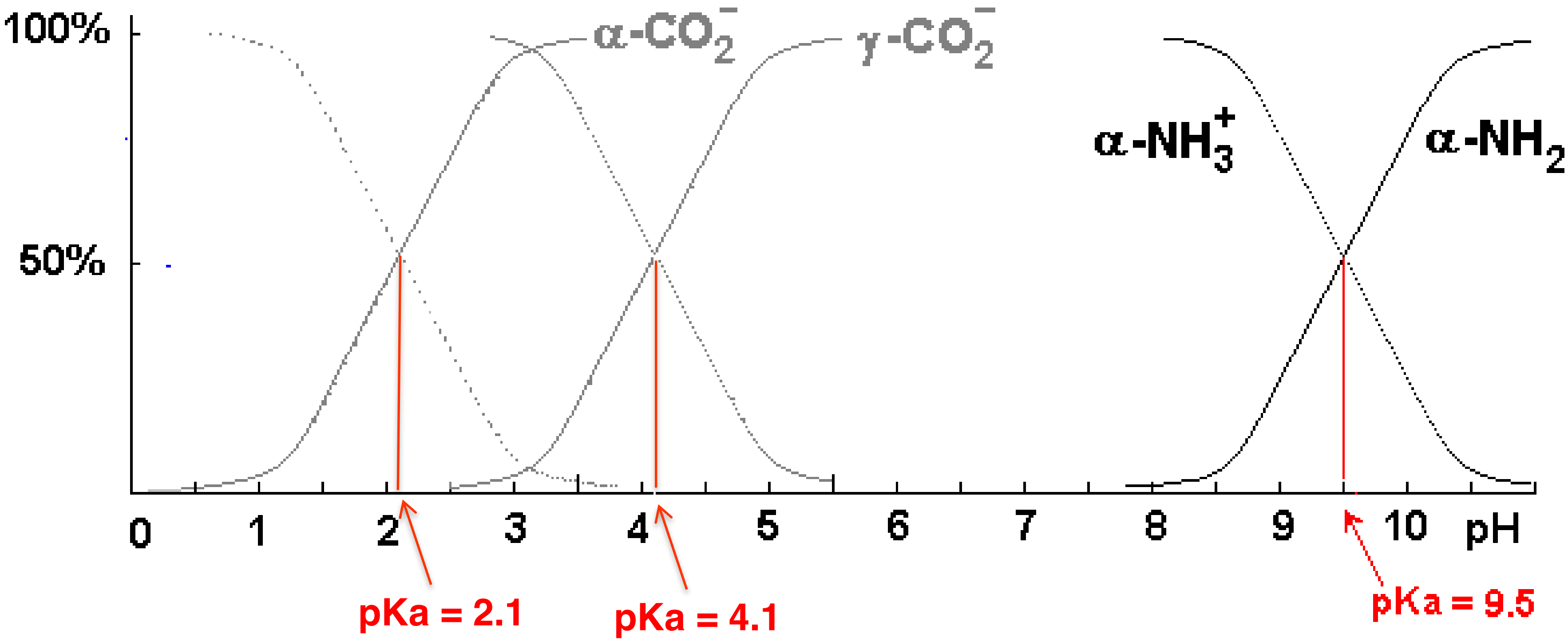
The value of pK_a tells you <u>where</u> in the pH scale a group undergoes deprotonation

e.g. **glutamate**

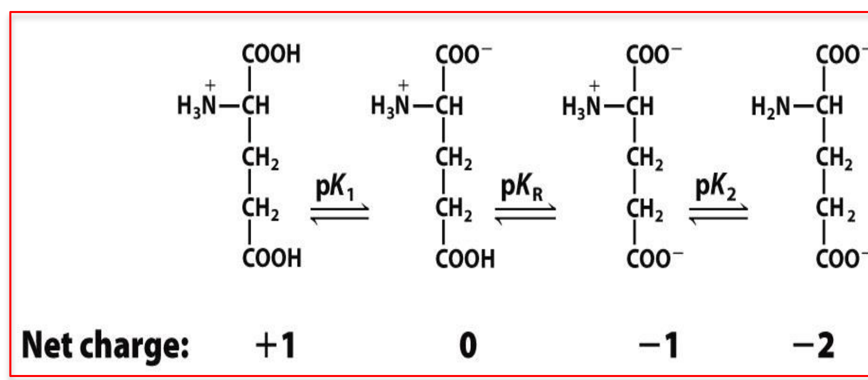




- A molecule can have several ionizable groups
- Each group has its own pK_a value
- Value of a pK_a depends on its chemical context
- an amino acid will have a slightly different pK_a when it is part of a peptide chain



Starting with glutamic acid at very low pH, all three functional groups are fully protonated. As we raise the pH, $[H^+]$ becomes less available, so deprotonation is more likely to occur.



Each group undergoes a transition as pH shifts from 0 to 14, starting ~1 unit below its pK_a and is almost complete by ~1 unit above its pK_a

How to assess the state of ionization of a functional group

- If **pH** is one unit or more **higher than pK_a**, the group is fully deprotonated

  Carboxylate groups with pK_a=2.4 exist as $–COO^-$ not –COOH at pH 7

- If **pH is equal to pK_a**, the group is 50% deprotonated and 50% protonated

- If pH is one unit or more **below the pK_a**, the group is fully protonated

  Amino groups with pK_a = 9.6 exist as $-NH_3^+$, not $-NH_2$ at pH 7

If pH is less than one unit away from pK_a, a calculation may be needed to determine the exact state

The ionization state of each group depends on its pKa value and the pH of the solution.

	pK_a*	pH 1 unit or more below the pK_a (protonated)	pH 1unit or more above the pK_a (deprotonated)	State at pH 7
Aspartate	4.0	-COOH (Neutral)	$-COO^-$ (-1)	$-COO^-$ (-1)
Glutamate	5.0	-COOH (Neutral)	$-COO^-$ (-1)	$-COO^-$ (-1)
Histidine	6.5	NH^+ (+1)	ring N: (Neutral)	76% ring N: (Neutral) 24% NH^+ (+1)
Lysine	10.2	$-NH_3^+$ (+1)	$-NH_2$ (Neutral)	$-NH_3^+$ (+1)
Arginine	12.5	$=NH_2^+$ (+1)	=NH (Neutral)	$=NH_2^+$ (+1)
Cysteine	8.5	-SH (Neutral)	$-S^-$ (-1)	–SH (Neutral)
Tyrosine	10.0	-OH (Neutral)	$-O^-$ (-1)	-OH (Neutral)
α-Amino	9.5 (9.6**)	$-NH_3^+$ (+1)	$-NH_2$ (Neutral)	$-NH_3^+$ (+1)
α-carboxylate	2.5 (2.4**)	-COOH (Neutral)	$-COO^-$ (-1)	$-COO^-$ (-1)

*pK_a values for amino acids in a polypeptide. (The charge associated with the side chain at each ionization state is shown within brackets) *(do not have to memorize pK_a values)*

**in a free amino acid

- The relationship of pH and pK_a tells you whether a group is __protonated____ or __deprotonated___, **NOT whether it is positive or negative**

- Groups that ionize on **O or S** atoms are **neutral when protonated,** and **negative when deprotonated**

$$-OH \rightleftharpoons -O^- + H^+ \qquad -SH \rightleftharpoons -S^- + H^+$$

- Groups that ionize on **N** are **positive when protonated,** and **neutral when deprotonated**

$$-NH_3^+ \rightleftharpoons -NH_2 + H^+$$

- ***There is no group that goes from positive to negative when it becomes deprotonated!!***

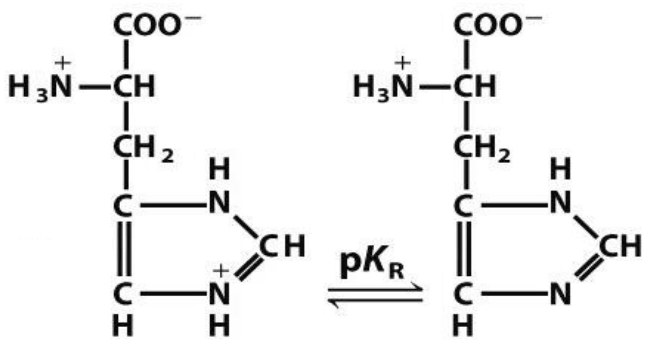
Calculating the exact state of ionization of a group at a given pH

- Histidine side chain has $pK_a = 6.5$
- At pH = 7, the major form will be deprotonated His, but some $HisH^+$ is present

$$pH = pK_a + \log\frac{[His]}{[HisH^+]}$$

$$7.0 = 6.5 + \log\frac{[His]}{[HisH^+]}$$

$$\frac{[His]}{[HisH^+]} = 10^{0.5} = 3.2$$



If ratio [His] / [HisH$^+$] = 3.2, what percentage of the total histidine is deprotonated?

We define α, **degree of deprotonation**, as the **fraction of histidine molecules that are deprotonated,** i.e. [His]/[total]

The **fraction protonated** is $1 - \alpha$

The **ratio** [deprotonated] / [protonated] = $\dfrac{\alpha}{1-\alpha} = 3.2$

$$\alpha = 3.2 - 3.2\alpha$$

$$\alpha\,(1 + 3.2) = 3.2$$

$$\alpha = \frac{3.2}{1+3.2} = 0.76$$

At pH 7, histidine is **76% deprotonated** and **24% protonated**

Molecules **exchange H$^+$** millions of times per second

A given His molecule is protonated 24% of the time

Averaged over time, charge on His at pH 7 is

(0.24 x +1) + (0.76 x 0) = **+0.24**

Home work questions

1. What is the overall charge on the side chain of lysine when one third of the side chains are deprotonated?

   A. +0.67

   B. +0.33

   C. -0.33

   D. -0.67

2. What is the overall charge on the side chain of aspartate when one quarter of the side chains are deprotonated?

   A. +0.75

   B. +0.25

   C. -0.25

   D. -0.75

3. Calculate the net charge of the **side chain** of Arg at pH 11.8. (Arg pK_a = 12.5)

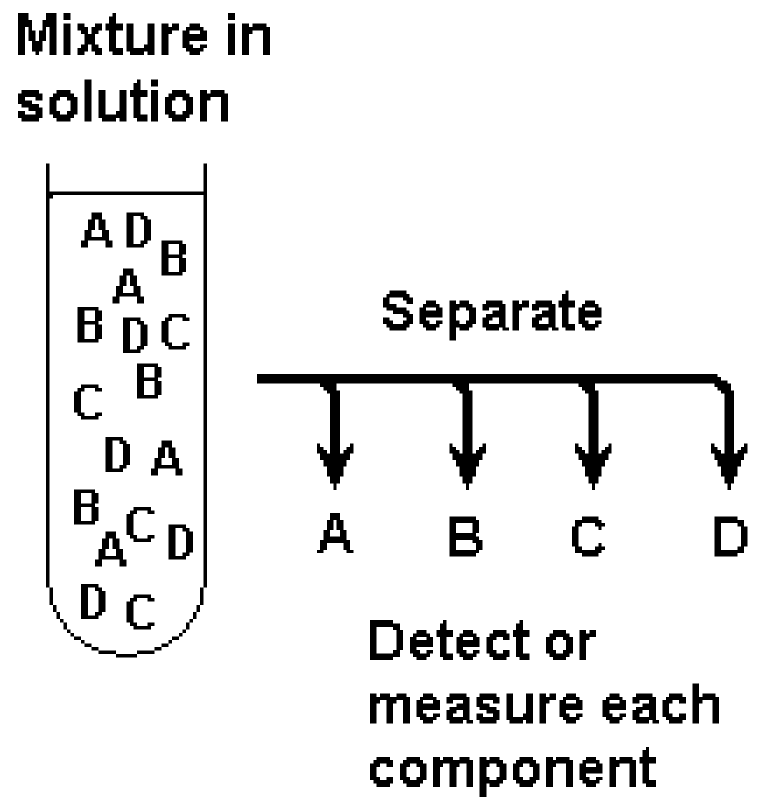
Answers

1. A
2. C
3. +0.83

Amino acid analysis

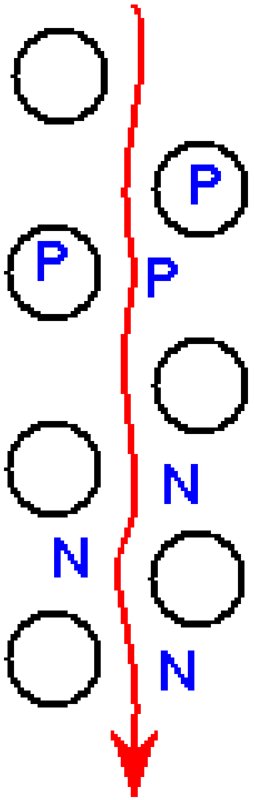
- Amino acid analysis helps to determine protein structure
- Analysis involves two processes:

1. Separation of a mixture into components
2. Detection of the components of interest
    - can be **qualitative** (tells you what is present)
    - can be **quantitative** (tells you how much is present)
    - can be **<u>preparative</u>** (separated components can be recovered for further experiments)



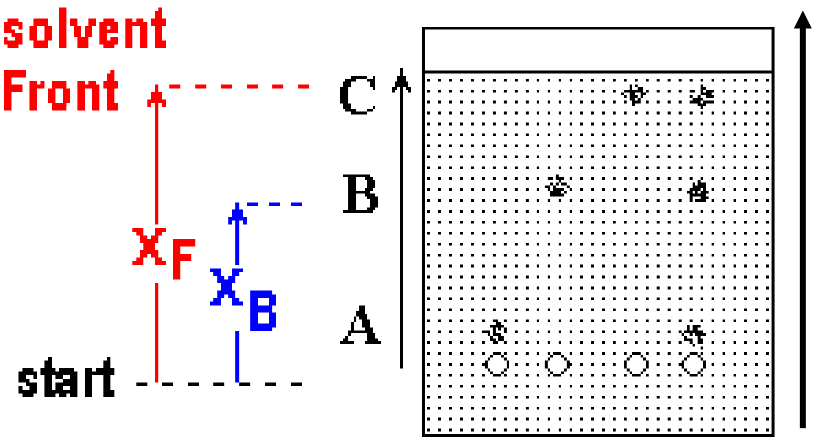
Partition chromatography is an important method for separating components of a mixture

- Particles of solid are chosen with a specific property, e.g. silica gel has **HO-Si-OH** groups that can hydrogen-bond to polar amino acids
  - Stationary phase
- Liquid solvent or buffer flows past the particles and is non-polar
  - Mobile phase
- **Amino acids exchange (partition) between phases**
  - polar amino acids **P** spend more of their time hydrogen bonded to silica and move **slowly**
  - non-polar amino acids **N** spend more time in solvent, and move almost as fast as solvent



Thin layer chromatography

- Silica gel is spread in a thin layer on a plastic sheet
- Samples are applied near the lower edge
- The lower edge is placed in solvent
- As solvent soaks up the sheet, different components of sample move with the solvent at different rates



Relative mobility

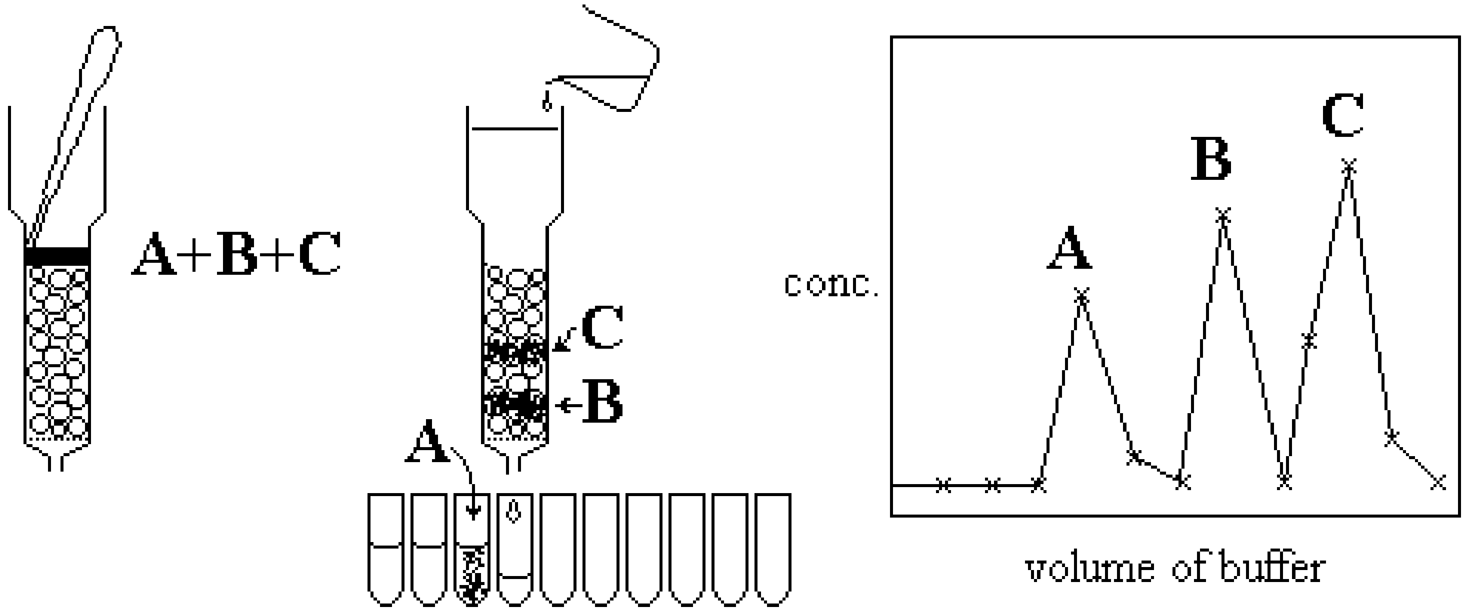
$$R_F = \frac{X_B}{X_F}$$

solvent has run to the top
A, B, and C have separated

The highest point reached by solvent is the __solvent front__

Each amino acid can be identified by its characteristic __relative mobility__ R_F
Very polar amino acids have low R_F, non-polar amino acids have high R_F

Another common format: column chromatography



1) Sample mix ABC applied at top

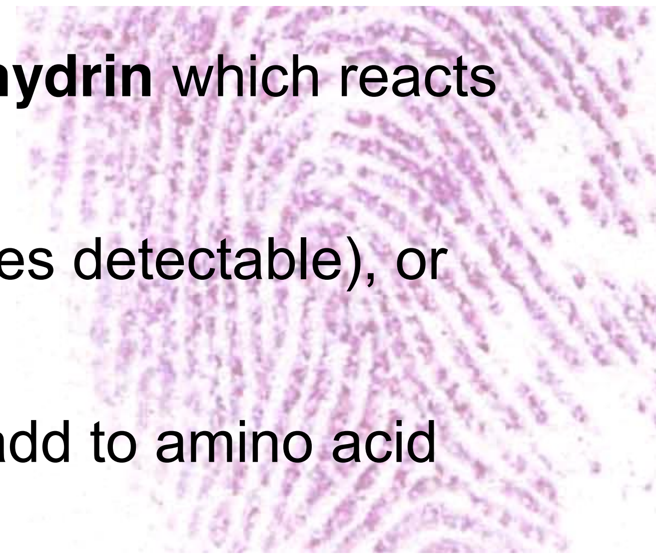
2) Buffer added, carries sample mix through column to collection tubes.

3) Contents of collection tubes analysed, results plotted.

Volume of buffer needed to move a compound through the column is the **elution volume.** Compounds can be **identified** by their characteristic elution volume

How are amino acids detected?

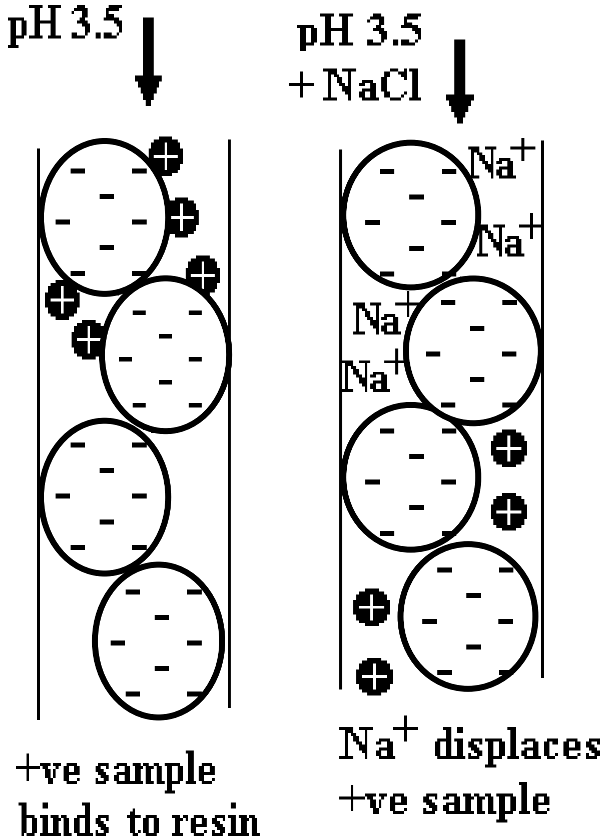
- Amino acids are colourless, and samples may be 10^{-6} to 10^{-10} moles
- They can be detected by adding **ninhydrin** which reacts with primary and secondary amines
- Gives intense purple colour (10^{-8} moles detectable), or yellow colour for proline
- Spray ninhydrin onto TLC plates, or add to amino acid solution, and heat
- Colour intensity is proportional to quantity of amino acid, and can be measured
- Alternative is **fluorescamine**, giving yellow fluorescence under UV light (10^{-10} moles detectable)



Different types of chromatography

- **Ion exchange chromatography** separates on the basis of charge

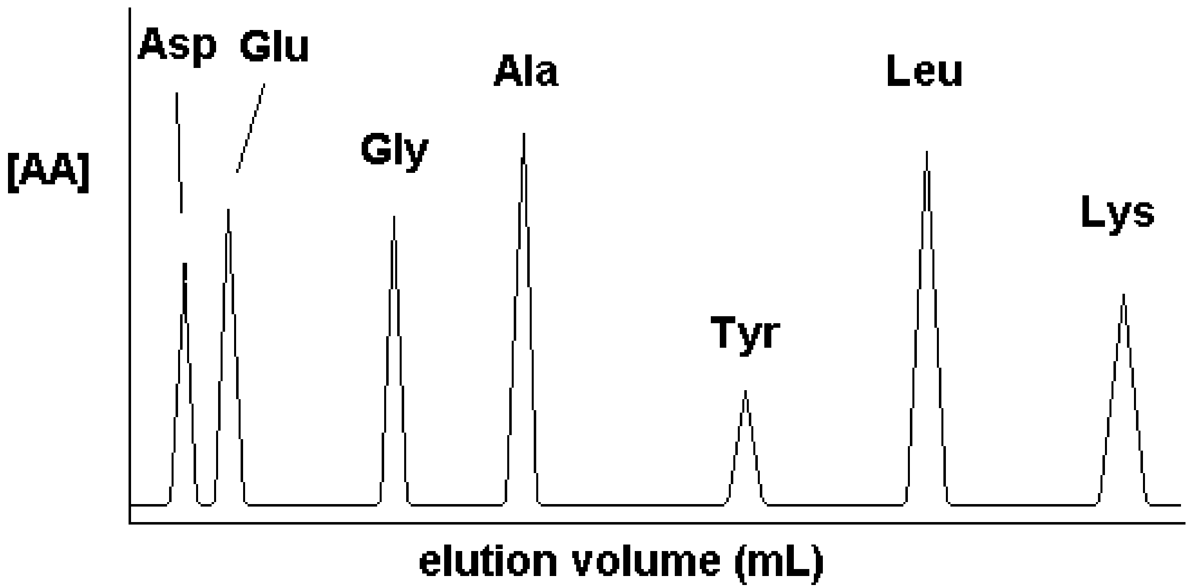
  - Uses **charged resins** as stationary phase
  - Cation exchanger resins contain negative groups, which bind **positive** molecules (cations) - Cation Exchange Chromatography
  - Anion exchanger resins contain positive groups, which bind **negative** molecules (anions) – Anion Exchange Chromatography
  - Elution is by:
    - Competition with a high ion concentration (usually NaCl), which displaces the amino acid from the resin
    - Changing the pH to alter the charge on the amino acid, so it no longer binds to the resin



- At pH 3.5, α-amino groups exist as $\mathbf{NH_3^+}$ while α-carboxylate groups exist mainly as $\mathbf{COO^-}$
- Side chains can also contribute to the charge
- The exact value of overall charge depends on **specific pKa values of the various groups in each amino acid**
- **Size of net charge** determines **how tightly** each amino acid binds
- High Na^+ present in elution buffer first **displaces weakly bound amino acids**. As **[Na^+] is increased**, **more tightly bound amino acids** are progressively displaced
- Alternatively, pH may be increased to eliminate the positive charge on the amino acid, so it no longer binds to the resin

Separation of amino acids by ion exchange

Amino acids are detected and their concentration measured in buffer coming out of the column



The volume of buffer needed to move a given amino acid from the top to the bottom of the column is the elution volume

- Elution volumes are often compared relative to a common standard, such as Ala or Leu
- Elution volumes are characteristic for each amino acid, and allow them to be identified

Separation of proteins from complex mixtures

- Proteins are derived from natural sources such as **microbial cultures**, plants, or animal tissues such as **liver**
- Extracts may contain **thousands** of different proteins
- Separation by **ion exchange is** based on **charge differences among proteins**
  - depends on the **relative number of Asp + Glu (negative)** versus **His + Lys + Arg (positive)** in each protein, and on **pH**
  - ~65% of all proteins are negatively charged at pH 7

Charge differences among peptides and proteins

e.g. at pH 7

Ala-Asp-Leu-Gly-His-Gln-Tyr-Cys-Ile-Glu-Lys-Ser-Thr

–1 **0.24** **–1** **+1**

due to side chains

+1 ------------ due to N- and C- terminal ---------------- **–1**

Net charge = +1 – 1 +0.24 – 1 + 1 – 1 = –0.76

midterm question: amino acids that are ionizable, remember N and C terminus on each end.

- Peptides and proteins can show large differences in charge
- Ion exchange is frequently used to separate protein mixtures

pKa values chart:

Side chain pKa values of amino acids:

Asp 4.0	Glu 5.0	His 6.5
Cys 8.5	Tyr 10.0	Lys 10.2
Arg 12.5		

α-amino group = 9.6
α-carboxylate group = 2.4

ex:

9.5	X	5.0	12.5	X	6.5	X	2.5
N -	Val -	Glu -	Arg -	Ser -	His -	Phe -	C
+1			+1		+1		-1

pH = 4.6

Average N-terminal amino group has pKa = 9.5
Average C-terminal carboxylic acid group has pKa = 2.5

---

pK_a values will be given at the exam. You are not required to memorize them.

Home work questions

1. Calculate the overall net charge of the following polypeptide at pH = 12.8.

   Ile-Gln-His-Arg-Trp-Asp

2. Which of the following peptides (represented by their one-letter codes) would bind most tightly to a cation-exchange column, at neutral pH?

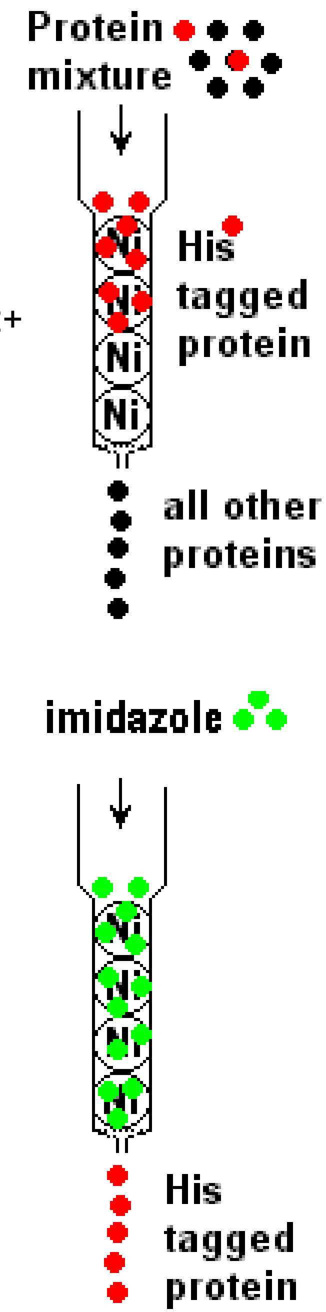
   A. CAMILLA
   B. EDWARD
   C. HARRY
   D. PHILIP
   E. WILLIAM

Answers

1. -1.67
2. C

Metal affinity chromatography

- Clusters of **His** in a protein bind tightly to **Ni^{2+} or Co^{2+}**
- Column is made up of _chelating resin_ **containing** Ni^{2+}
- If protein is artificially produced by inserting its gene into cells, the gene can be modified to include **6-8 extra His residues** at N- or C-terminus
  - The added His cluster is called a _His-tag_.
  - His-tagged proteins **binds tightly** to the Ni^{2+} resin
- The His-tagged protein is eluted by **adding** _imidazole_ (structure similar to His side chain) to the buffer
- Imidazole out-competes His-tag, and protein no longer binds to the column
- **High degree of purification in one step**



Separating proteins on the basis of size

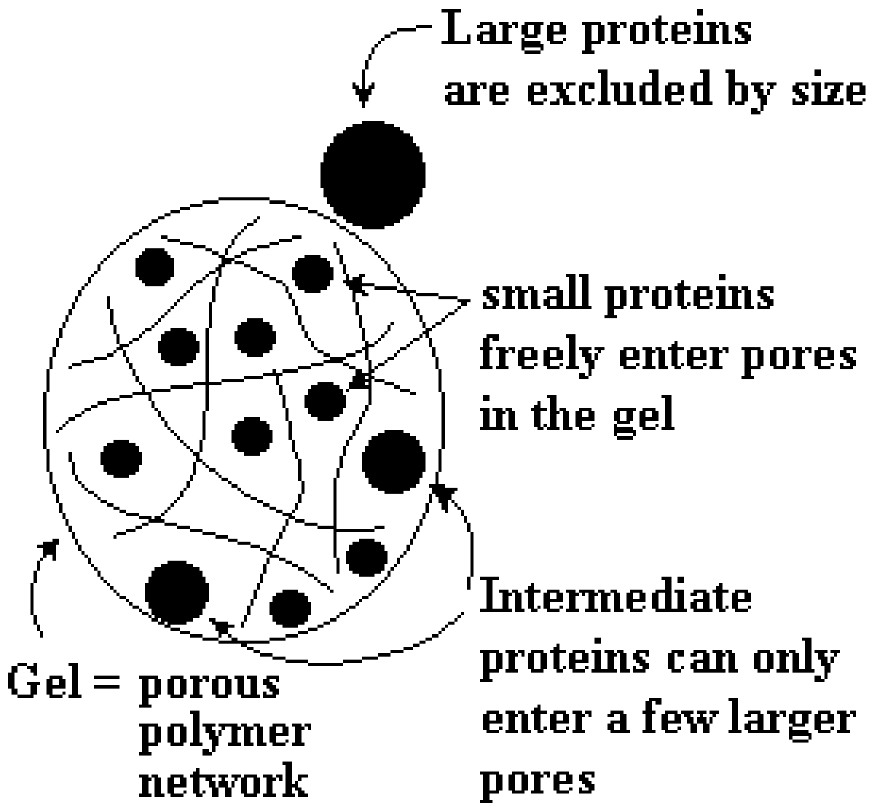
- Different proteins or peptides can vary widely in size (number of amino acids)

- **Gel filtration** or **molecular exclusion** chromatography allows separation of proteins **on the basis of** __size______**.**

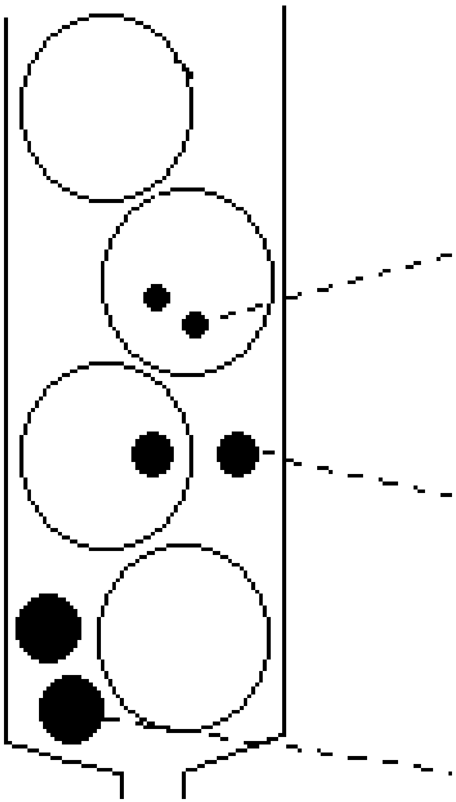
- Gel filtration can also be used to **measure the size** of an unknown protein

  Sample is compared to proteins of known size

- Beads of polymeric gel - a loose network of polymer with **many water-filled pores**.
- Protein molecules can enter the pores if they fit
- Larger proteins are excluded from the pores
- Proteins of intermediate size may enter some of the pores

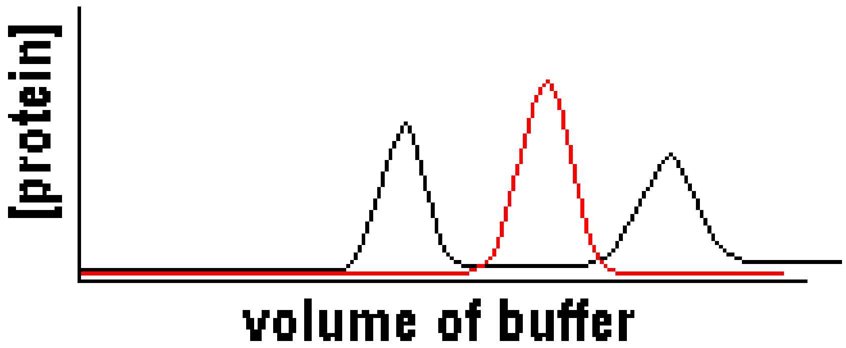


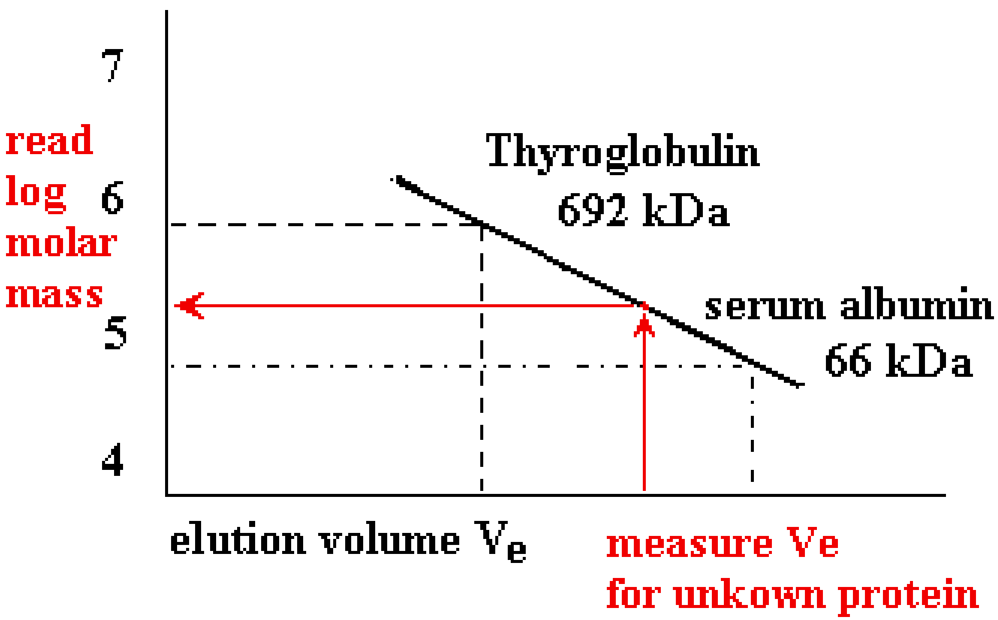
Column chromatography using gel filtration



- Small molecules enter pores and are delayed in their movement down the column
- Intermediate size molecules can only enter some pores and are delayed less
- Very large molecules are **excluded** from gel and stay in the buffer flowing around the beads – pass through the column quickly

Gel filtration can be used to measure the molar mass of proteins





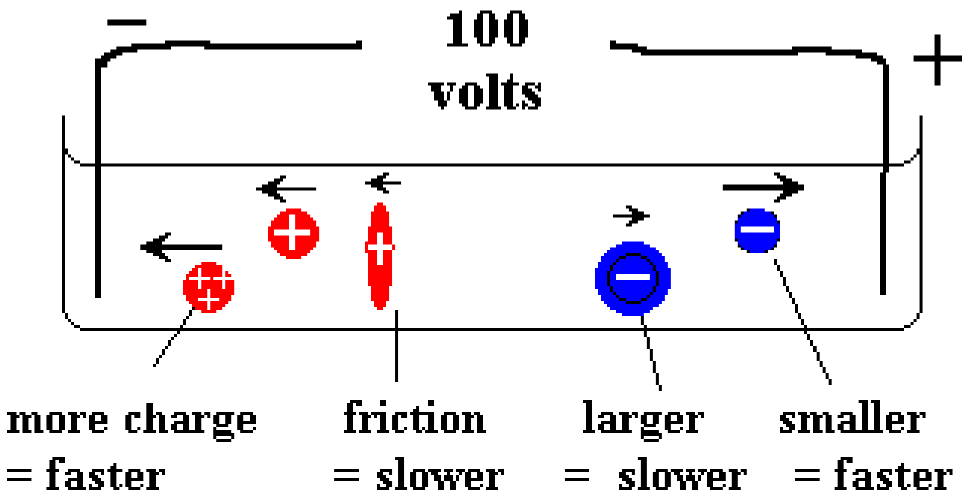
- Measure **elution volume** of proteins of known mass
- Elution volume $\mathbf{V_e}$ is the volume of buffer needed to **move a protein from top to bottom** of column
- **Elution volume is a linear function of log molar mass** (negative slope)
- Then **measure elution volume** of unknown protein and **project back to the log mass axis**

More methods for separating proteins

Ultracentrifugation:

- a protein sample is placed in an **ultracentrifuge** spinning at 10,000 to 75,000 rpm, producing a force of **10,000 to 500,000 x gravity (g)**
- At these g-forces, molecules **sediment** (move down the tube) at a rate that depends on **size and shape**
- By measuring _sedimentation velocity_, it's possible to calculate the molecular mass of a protein

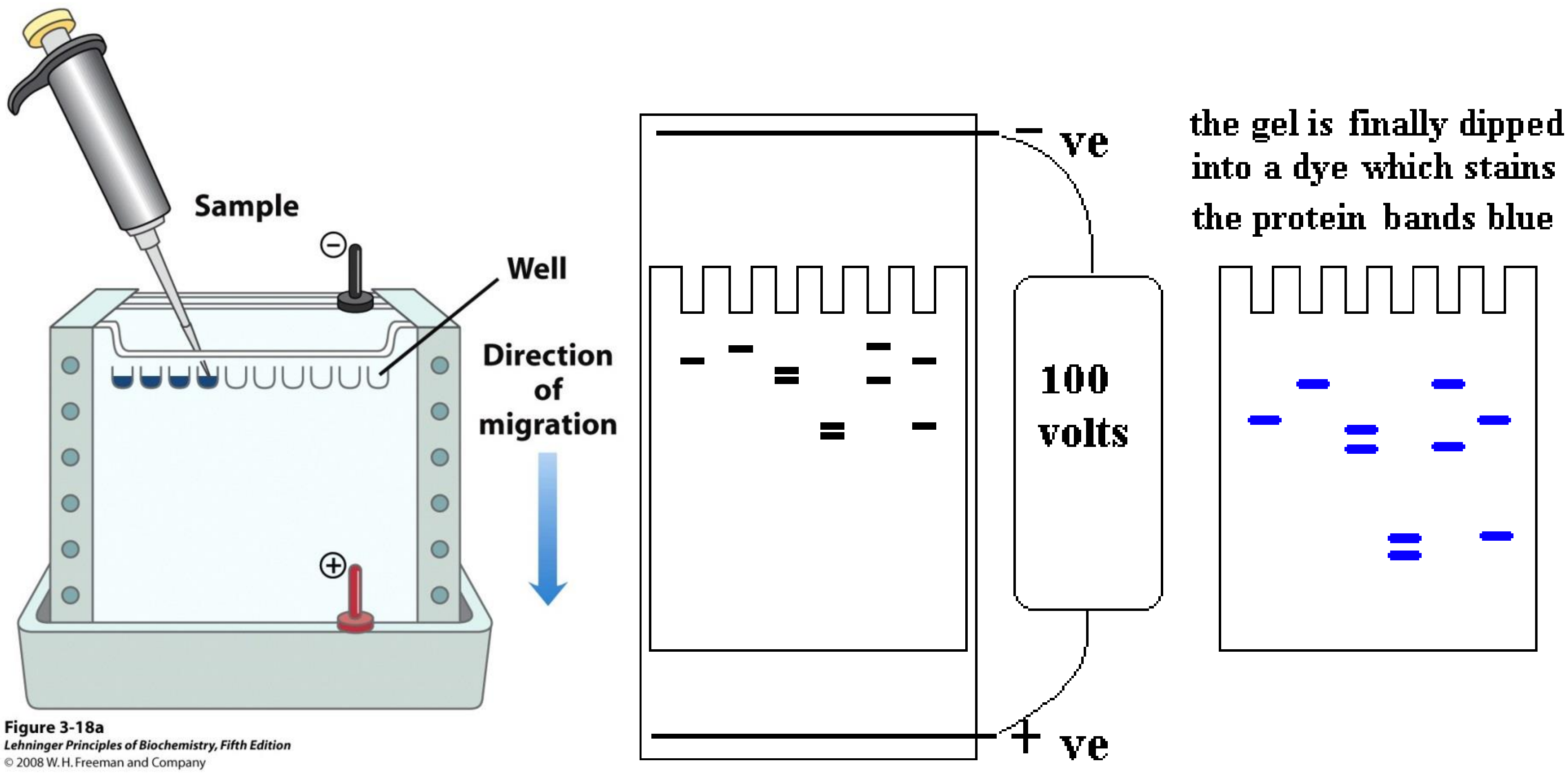
Electrophoresis: separation based on movement of charged molecules in an electric field



- Rate of movement depends on **size, _shape_ and charge**

- To prevent movement of the solution, the separation is carried out in a porous gel

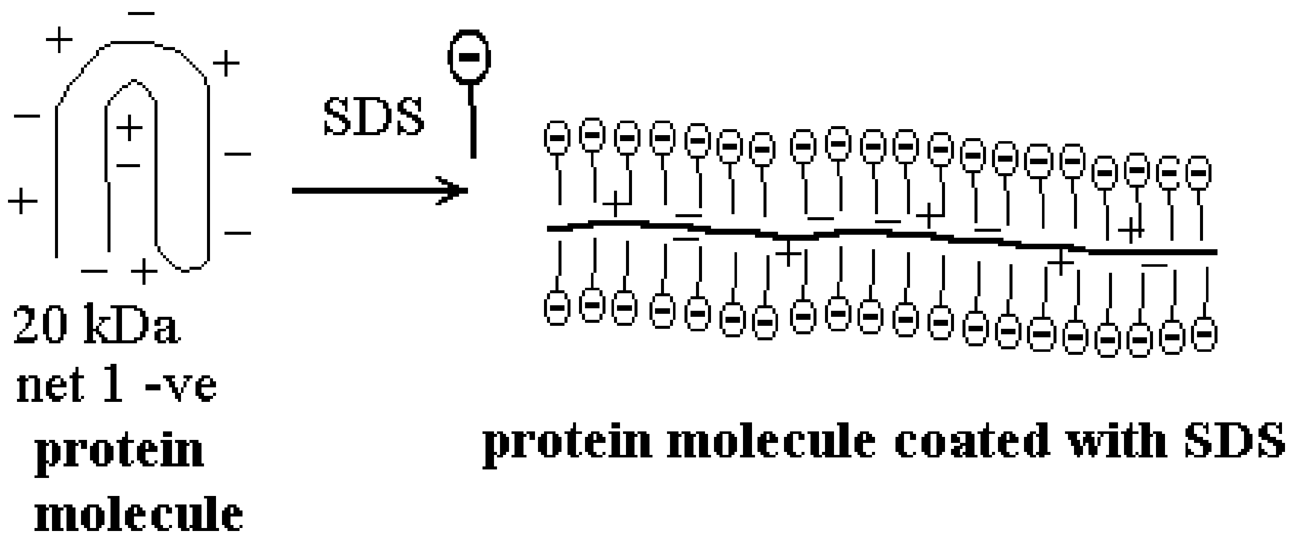
Polyacrylamide gel is easily prepared in lab, has right porosity for proteins 10 kDa -1000 kDa



- A typical gel is 5-15% polymer and 90-95% water with conductive buffer

In SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE) the protein is pre-treated with the detergent sodium dodecyl sulfate (SDS)

SDS causes protein molecules to **extend**, and gives a **uniform charge per unit size**

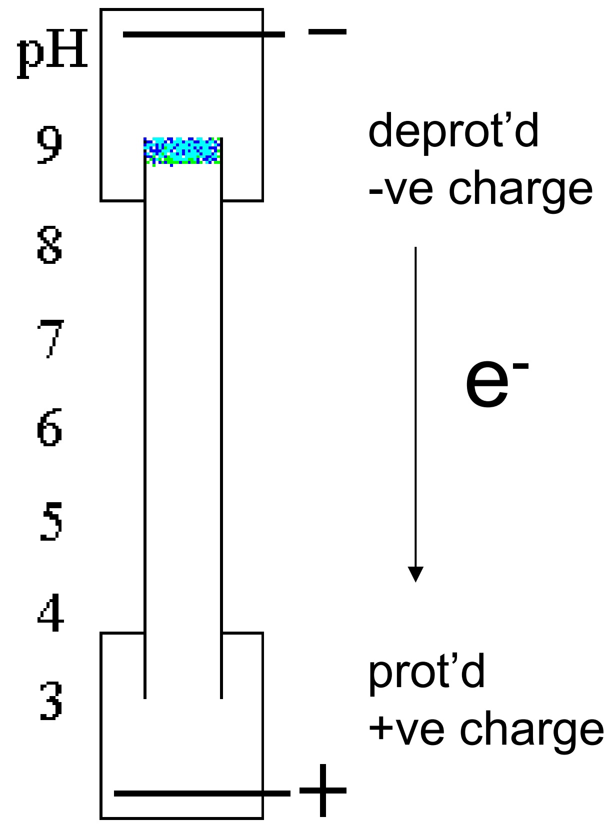


- When SDS is used, native protein charge is swamped out, and the overall –ve charge is used to move the protein through the gel (towards the + electrode)
- **separation is based strictly on** __size__
- small proteins can fit through all the pores, move rapidly down the gel
- large proteins have to "meander" to find pores to fit through, move more slowly
- SDS-PAGE may also be used to measure the molecular mass of a polypeptide chain by comparing with __standards__ of known size

Isoelectric focussing: separation based on charge

Isoelectric point, pI: pH at which the net charge on a protein is __zero__

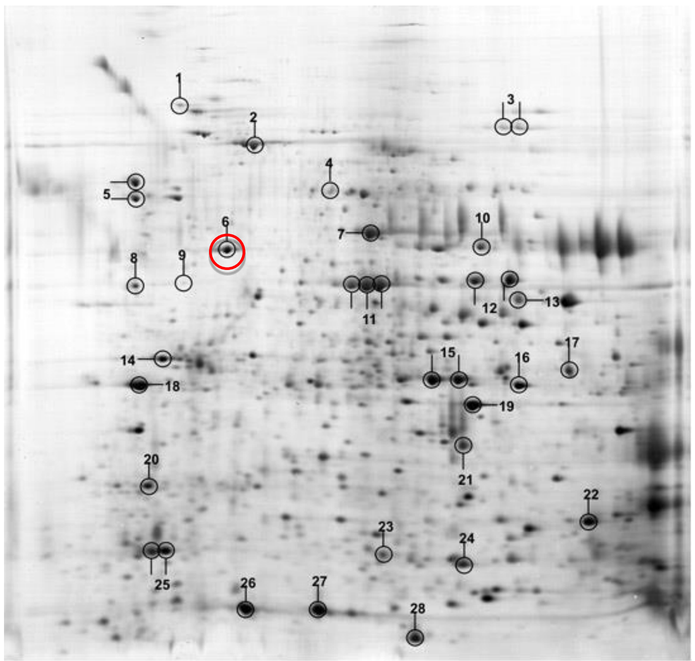
- at high pH, protein is deprotonated, moves toward the + electrode
- as it passes through gradient of decreasing pH, becomes protonated, net -ve charge decreases
- when the net charge = 0, protein stops moving
- each protein in a mixture has a different isoelectric point, so they become separated along the pH gradient



tube of gel with pH gradient from one end to the other

Two-dimensional gels: Separation of complex samples

- Combines isoelectric focusing and SDS electrophoresis
- Can separate individual proteins in very complex mixtures



Soluble proteins of *Schistosoma mansoni*

1 HSP86
2 HSP70
3 ATP:Guanidino Kinase
4 Adenylate Dehydrogenase
5 Calreticulin
6 Actin
7 Enolase
8 Tropomyosin
9 Serpin-like
10 Phosphoglycerate kinase
11 p40
12 Aldolase
13 GAPDH
14 14-3-3 e
15 GST28
16 Triose Phosphate Isomerase
17 Elongation Factor 1a
18 14-3-3 homolog 1
19 GST26
20 Calpain
21 Myosin Light Chain
22 Cycophilin
23 Superoxide Dismutase
24 Fatty Acid Binding Protein (Sm14)
25 SME16
26 Thioredoxin
27 Dynein Light Chain
28 Ubiquitin
29 Adenylate Kinase

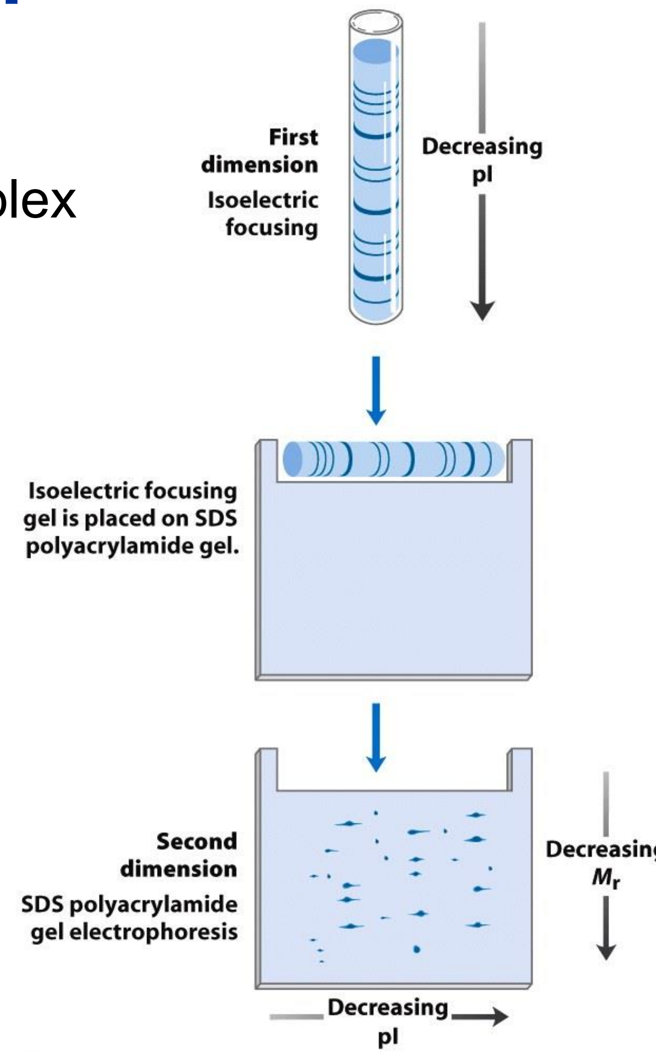
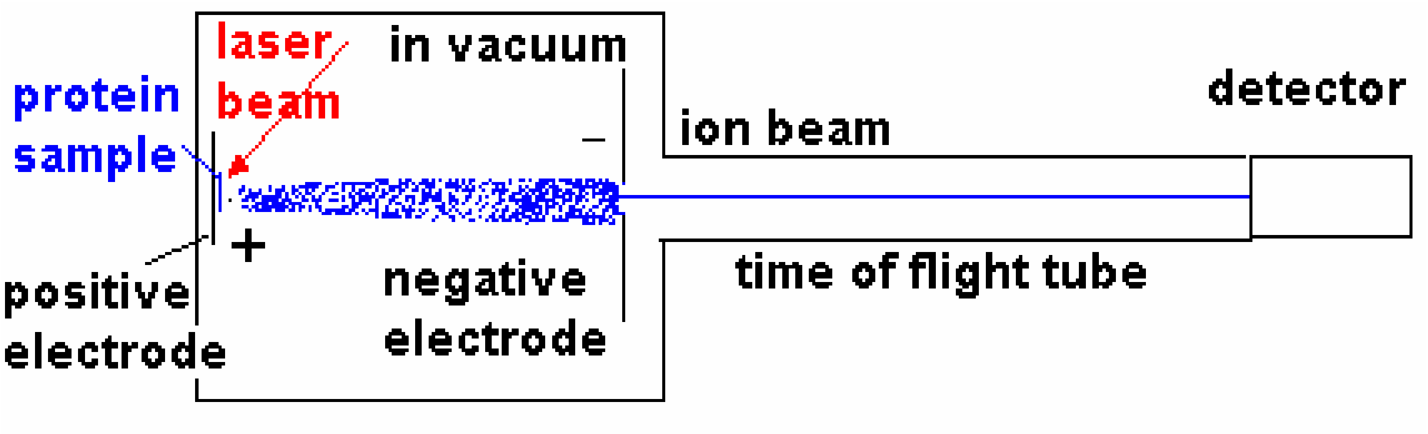


Figure 3-21a
Lehninger Principles of Biochemistry, Fifth Edition

Mass spectrometry provides a way to identify proteins



- A protein is vaporized by laser beam, yielding charged protein particles
- Particles travel toward the detector
- Velocity depends inversely on the mass of the particle (larger = slower)
- The __time of flight__ to the detector yields a very **accurate** mass measurement (5 significant figures of accuracy)
- We can compare the mass of the protein with a __database__ of proteins of known mass, allows identification

Polypeptides and proteins: structural hierarchy, sequence. Basis of reactivity and hydrolysis

A protein is a long linear chain of amino acids

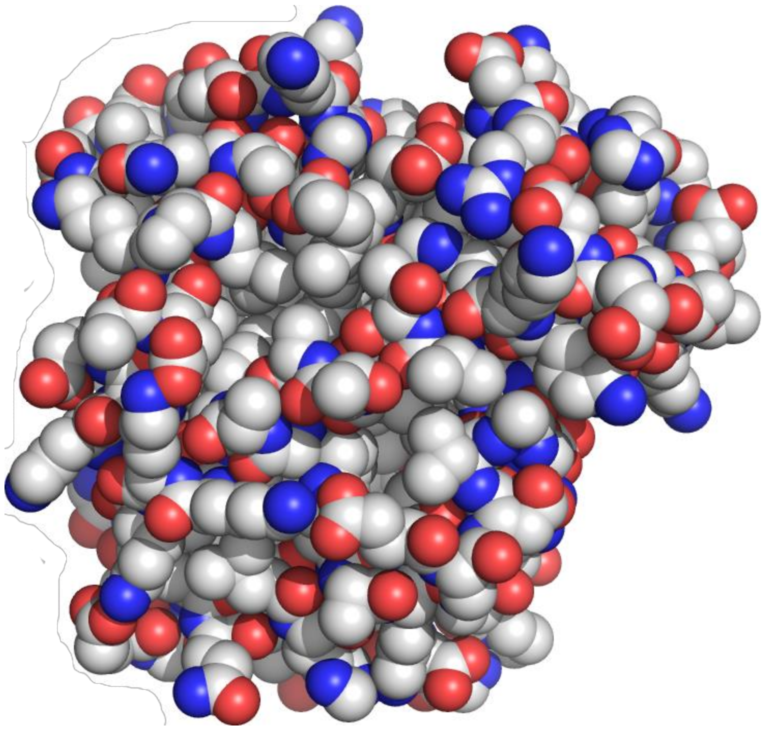
VAL	LEU	SER	GLU	GLY	GLU	TRP	GLN	LEU	VAL	10
LEU	HIS	VAL	TRP	ALA	LYS	VAL	GLU	ALA	ASP	20
VAL	ALA	GLY	HIS	GLY	GLN	ASP	ILE	LEU	ILE	30
ARG	LEU	PHE	LYS	SER	HIS	PRO	GLU	THR	LEU	40
GLU	LYS	PHE	ASP	ARG	PHE	LYS	HIS	LEU	LYS	50
THR	GLU	ALA	GLU	MET	LYS	ALA	SER	GLU	ASP	60
LEU	LYS	LYS	HIS	GLY	VAL	THR	VAL	LEU	THR	70
ALA	LEU	GLY	ALA	ILE	LEU	LYS	LYS	LYS	GLY	80
HIS	HIS	GLU	ALA	GLU	LEU	LYS	PRO	LEU	ALA	90
GLN	SER	HIS	ALA	THR	LYS	HIS	LYS	ILE	PRO	100
ILE	LYS	TYR	LEU	GLU	PHE	ILE	SER	GLU	ALA	110
ILE	ILE	HIS	VAL	LEU	HIS	SER	ARG	HIS	PRO	120
GLY	ASP	PHE	GLY	ALA	ASP	ALA	GLN	GLY	ALA	130
MET	ASN	LYS	ALA	LEU	GLU	LEU	PHE	ARG	LYS	140
ASP	ILE	ALA	ALA	LYS	TYR	LYS	GLU	LEU	GLY	150
TYR	GLN	GLY								

Myoglobin

O_2 binding protein from muscle

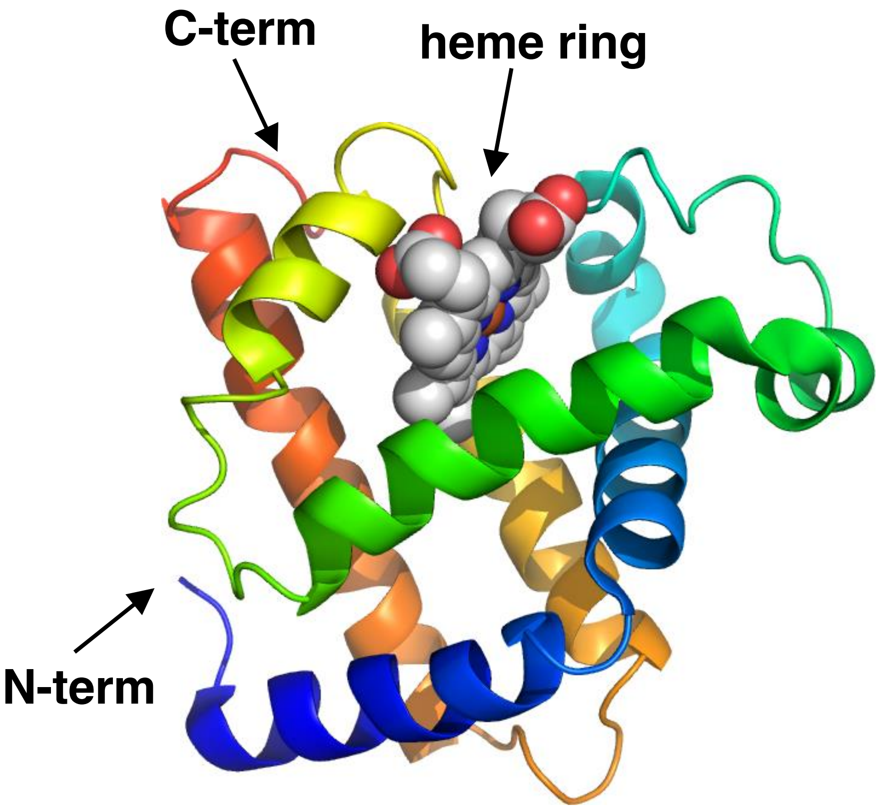
153 amino acids in chain

The linear chain of myoglobin is precisely folded into a compact, 3-dimensional globular structure



- This image is an accurate representation of the structure of myoglobin
- Too much detail makes it hard to comprehend

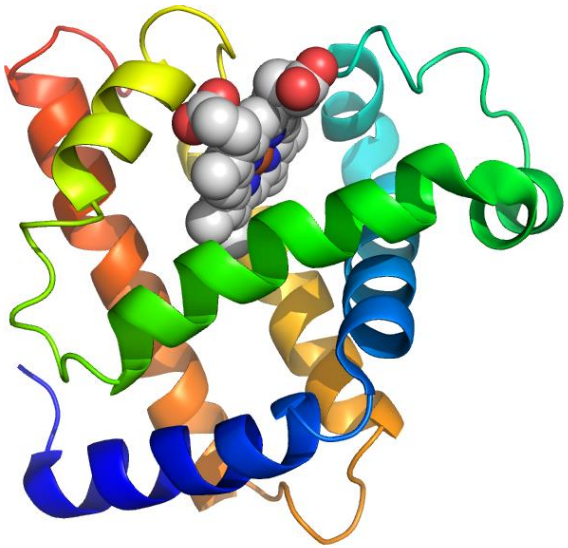
The linear chain of myoglobin is precisely folded into a compact, 3-dimensional globular structure



- A simplified view shows distinct regular patterns in folded protein
- The ribbon **traces the path of the** _polypeptide backbone_ (peptide bonds only, no side chains)
- Colour coding: blue at N-terminal progressing to red at C-terminal
- In myoglobin, the ribbon is arranged in eight spiral or **helical** segments
- Amino acid side chains fill the “spaces”, which are not empty

Protein structure is organized in three (or four) levels

- **Primary structure**: the linear **sequence** of amino acids
- **Secondary structure**: regular repetitive patterns, such as the **helical sections** in myoglobin, that run along short sections of peptide chain (5-20 AAs)
- **Tertiary structure**: the **overall pattern of 3D folding** of the whole polypeptide
  - in myoglobin, 8 helical sections enclose a central cavity
- Some proteins also have **a quaternary level of structure**
  - e.g. **hemoglobin is an assembly, or complex, of four globin units,** each one similar to myoglobin
  - the four globins act **cooperatively** to improve oxygen transport



Protein function depends on positioning of key amino acids close to each other in 3-D space, even though they are far apart in the linear sequence.

The exact pattern of folding is **critical** for protein function.

Protein structure is organized in three (or four) levels

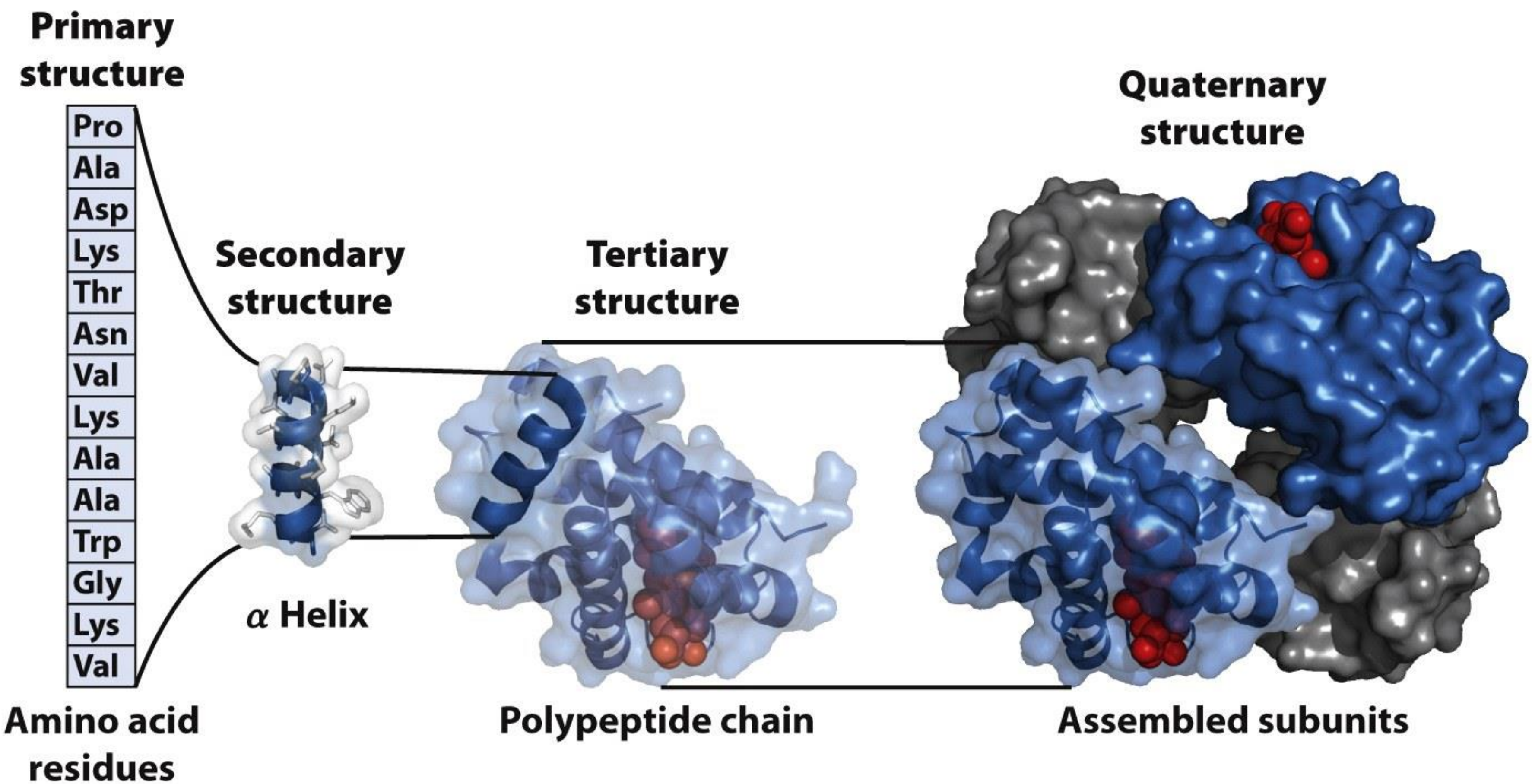


Figure 3-23
Lehninger Principles of Biochemistry, Fifth Edition

To investigate the structure of a protein, we find out which amino acids are present, and in what sequence

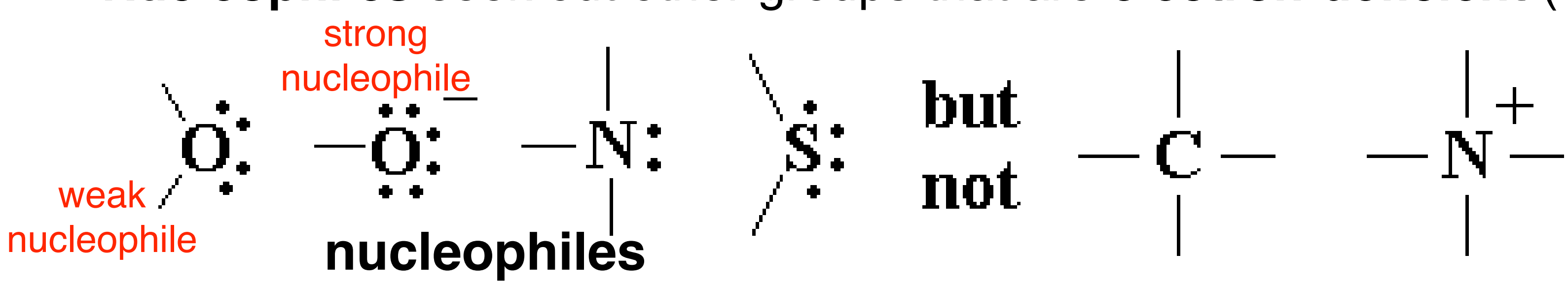
- Peptide bonds of protein are __hydrolysed__ with help of catalyst to release individual amino acids
  - Acid hydrolysis: **6 M HCl 110º**, 24-72 h to completion
    - but Trp is destroyed
  - Base hydrolysis: **4 M NaOH, 110º**, 16 h to completion
    - but some other amino acids (not Trp) may be destroyed
  - With digestive enzymes called __proteases__:
    - enzymes are proteins which have a catalytic function
    - proteases catalyse hydrolysis of peptide bonds
- After hydrolysis of the protein, amino acids can be analyzed by chromatography – tells you **how much** of each amino acid is present

Basis of chemical reactivity in hydrolysis and other biochemical reactions

- **Chemical reactivity** arises from an unbalanced distribution of valence electrons
  - C-C and C-H bonds **share bonding electrons equally** and are both **non-polar** and chemically **unreactive**

- Reactions involve atoms which have **unshared electrons (lone pairs),** or are **electron deficient,** or **pull electrons towards them** due to **electronegativity**

- Many biochemical reactions are **initiated by <u>nucleophiles</u>**
  - **A** _nucleophile_ ("nucleus-lover") **is an atom with a lone pair of electrons available to share**
  - **Nucleophiles** seek out other groups that are **electron-deficient** ("nuclei") -electrophiles



An atom with a lone pair may use it in different ways

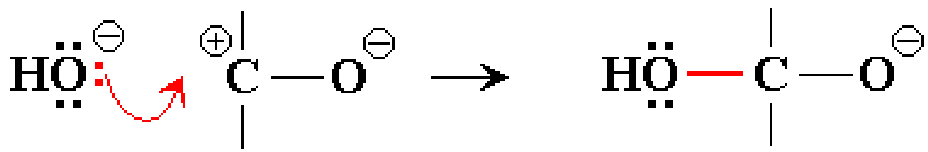
- As an H-bond acceptor if it simply attracts an **O-H** or **N-H**

  R-O: - - - H–N-R

- As a **base if it captures H^+**

  $\mathbf{H^+ + :NH_2\text{–}R \rightarrow {}^+NH_3\text{–}R}$

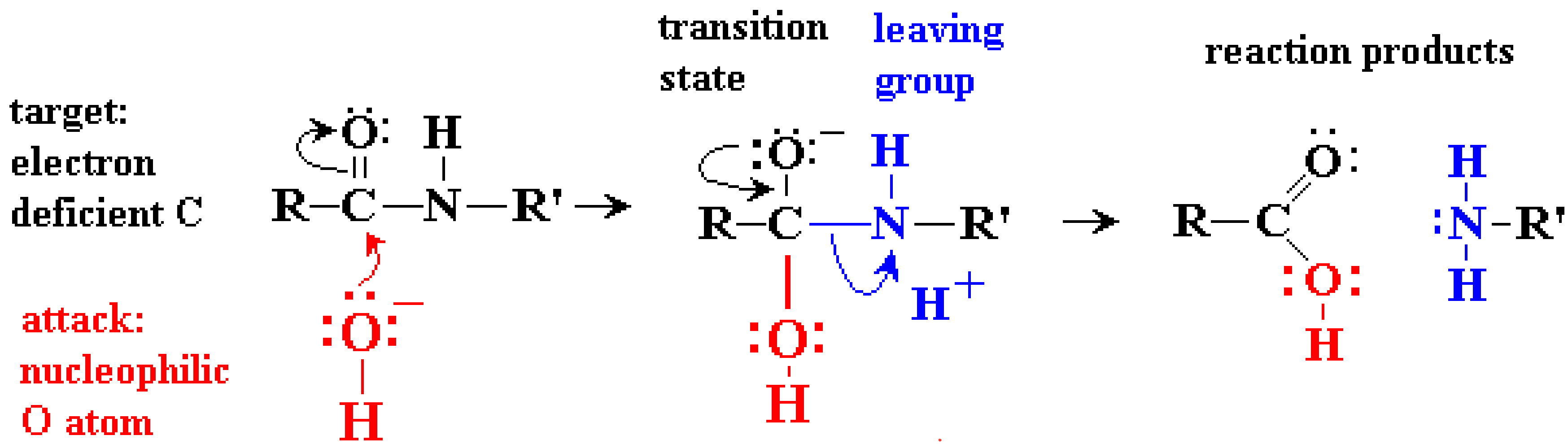
- As a nucleophile when it **shares** the lone pair with another electron-deficient atom to make a **new bond**



Nucleophilic substitution or displacement: incoming nucleophile X attacks **target atom C** to displace **leaving group Y**

X: C—Y → X—C :Y

Hydrolysis is attack by H_2O as nucleophile on the electron deficient C of the peptide bond



- **C is electron-deficient** (an **electrophile; electron-lover**), because the **electronegative O** bonded to C pulls electrons away from it
  - this lets C take up the **incoming electron pair** in a **new bond**
  - **Curly arrows** represent _movement of electron pairs_
  - The **transition state** is semi-stable, like a compressed spring
- **Amino N** acts as **leaving group**, allowing the peptide bond to break
- Nucleophilic reactions - see Lehninger p 216 (208)

Cut these strips out and bring to next class !

| Met Asn Lys |

| Asp Ile Ala Lys |

| Glu Ala Leu Phe Arg |

| Glu Leu Tyr Gln Gly |

| Gln Gly |

| Met Asn Lys Glu Ala Leu Phe |

| Arg Asp Ile Ala Lys Glu Leu Tyr |