

BMOL2201/6201

Biochemistry: An Introduction

Shoba Ranganathan

Applied Biosciences

T: 02 9850 6262; E: shoba.ranganathan@mq.edu.au

Important messages

Unit Information

Lecture Notes

Practicals and Tutorials

Assessment Outlines

Practice Quizzes

- iLearn.mq.edu.au has all the info you need for this unit.
- Course delivery:
 - 2 lectures each week (recorded)
 - 5 x 3 hr pracs alternating with 5 x 2 hr tutorial (or SGTA): Wed pm, Thu am and Thu pm.
 - Your class number is the same for Prac and SGTA as they run on alternate weeks, starting on the same day at the same time.
- Classes 1, 2 & 3 start with Labs from Academic Week 2 (on 3 & 4 March)
- Classes 4, 5, & 6 start with SGTA from Academic Week 2 (on 3 & 4 March)
- **This week: pl. go to the LABS and complete your safety form after watching the safety video, as per the schedule below:**
 - **Wed pm: Class 1: 2 - 2.30 pm & Class 4: 2.30-3 pm**
 - **Thu am: Class 2: 10-10.30 am & Class 5: 10.30-11 am**
 - **Thu pm: Class 3: 2 - 2.30 pm & Class 6: 2.30-3 pm**

What is Biochemistry ?

- **The study of life: chemistry of biomolecules**

➤ overlaps other disciplines, including cell and molecular biology, chemistry, genetics, immunology, microbiology, pharmacology, and physiology

- **Key questions**

1. What are the structures of biological molecules?
2. How do biological molecules interact with each other?
3. How does the cell synthesise and degrade biological molecules?
4. How is energy conserved and used by the cell?
5. What are the mechanisms for organizing biological molecules and coordinating their activities?
6. *How is genetic information stored and processed?*

BMOL2201/6201

- Building blocks of key biomolecules
 - ❖ nucleic acids,
 - ❖ proteins,
 - ❖ sugars and
 - ❖ lipids
- Biochemical signalling: how does the cell know what to do and when?
- Metabolism: making biomolecules from their building blocks as well as breaking them down
- How do we get energy?
- How are biomolecules organized into teams for coordinating their activities?

Eat healthy and avoid metabolic diseases such as obesity and diabetes!

Learning Objectives of this unit

1. Define the structural and metabolic differences between eukaryotic and prokaryotic cells with emphasis on biochemical energy metabolism, involving the synthesis and breakdown of important biomolecules.
 - identify the differences in cell structure;
 - relate the biochemical processes required for growth and energy; and
 - understand control mechanisms involved in the metabolism of important biomolecules.
2. Define chemical and biochemical principles and apply these to identify the interactions between different metabolic pathways and the biochemical signals.
 - the key concepts of compartmentation of biochemical processes;
 - the major biological systems involved in metabolism and energy production pathways in the living cell; and
 - understand how external changes are communicated to the interior of cells and organelles.

Learning Objectives of this unit - 2

3. Connect protein structure with function:
 - by defining the protein structure-function paradigm; and
 - evaluate the relationship between structure and function of proteins.
4. Identify, quantify and separate biomolecules using appropriate experimental methods:
 - utilize appropriate experimental methods to
 - ❖ characterise,
 - ❖ quantify and
 - ❖ separate different types of biomolecules.

Learning Objectives of this unit - 3

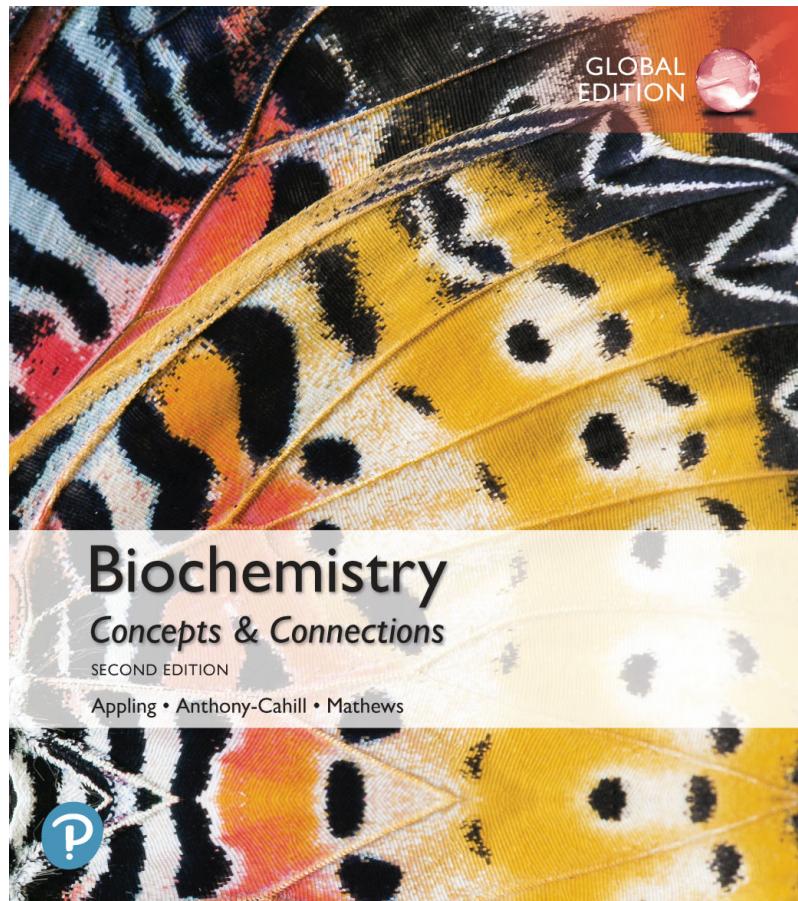
5. Track and measure rates of enzyme reactions and calculate kinetic parameters from the data generated:
 - understand the thermodynamic principles of enzyme catalysis;
 - identify the factors influencing the rate of an enzyme-catalysed reaction; and
 - measure rates of enzyme reactions to determine basic kinetic parameters of an enzyme.
6. Collect experimental data using biochemical techniques and sort, graph, analyze and present the experimental results in a biochemical context
 - utilize appropriate experimental methods to
 - ❖ characterise,
 - ❖ quantify and
 - ❖ separate different types of biomolecules.

These learning objectives will be delivered **via lectures, practical exercises** (in the laboratory) and **tutorial questions** (in SGTA sessions).

Content delivery

- **iLearn:** ilearn.mq.edu.au
 - Login: MQ Student ID (8 numbers)
 - Password: your MQ password (OneID)
 - ❖ echo360 lecture video recordings
 - ❖ Announcements and Discussion Forum
 - ❖ Prac and Tutorial materials
 - ❖ Link to textbook and assessments.
- **Email:**
 - All emails will be sent to your ***MQ student email***
 - Please setup email forwarding to your favourite email address if you do not regularly look at this

Textbook



Biochemistry: Concepts and Connections

Second Edition, Global Edition

Dean R. Appling
Spencer J. Anthony-Cahill
Christopher K. Mathews
(AAM)

- Link to Textbook resources from iLearn
 - Uses **Mastering Chemistry** adaptive learning approach
 - **eText**
 - Revision exercises
 - Practice exercises
 - Assessment tests
 - Online tutorials.



MACQUARIE
University

Module 1: Building Blocks of Biochemistry

Where is the e-text?



Student Links

Welcome to MasteringChemistry

Mastering Assignments

More... 

Student Links

Welcome to MasteringChemistry

Mastering Assignments

Mastering Scores (for students)

User Settings

Pearson eText

MyLab and Mastering Course Home

Study Area

Pearson Announcements

 Less

Module 1: Building Blocks of Biochemistry



MACQUARIE
University

Brushing up your fundamentals

- Revision materials on Textbook Resources
- Check out the review materials at the Pearson Mastering site.

Tips on doing well

- You'll need to work steadily, because new material is based on earlier material.
- You will not be able to leave mastering the content until the last few weeks.
- **Direct correlation between lecture/tutorial attendance and student performance.**

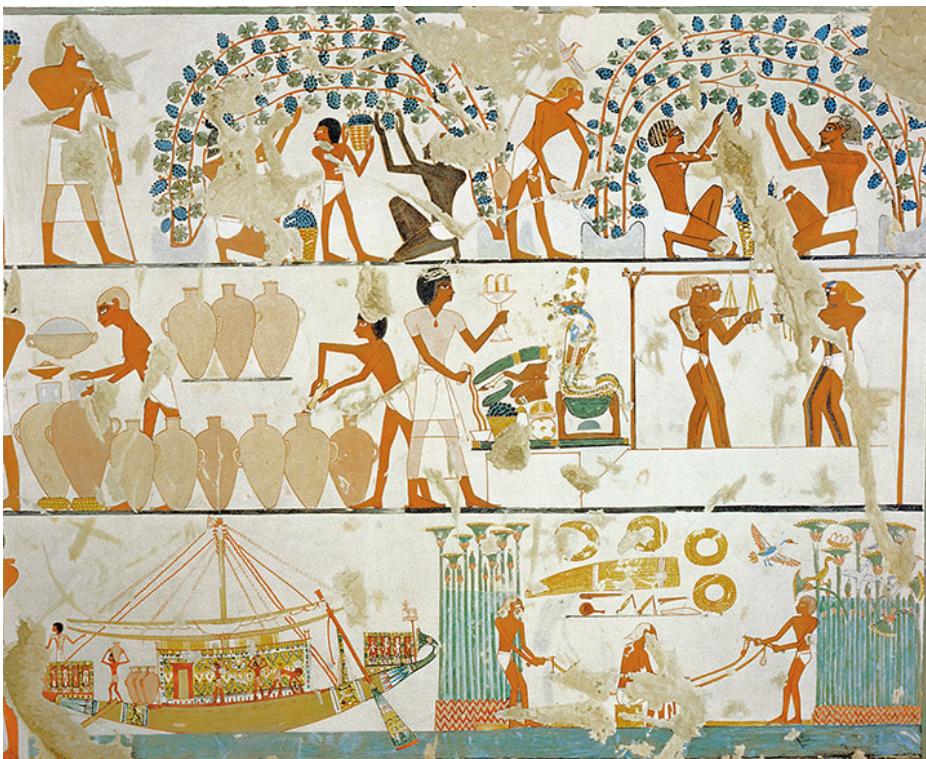
Biochemistry and the Language of Chemistry

- The Science of Biochemistry
- The Elements and Molecules of Living Systems
- Distinguishing Characteristics of Living Systems
- The Unit of Biological Organization: The Cell
- Biochemistry and the Information Explosion

AAM: Chapter 1

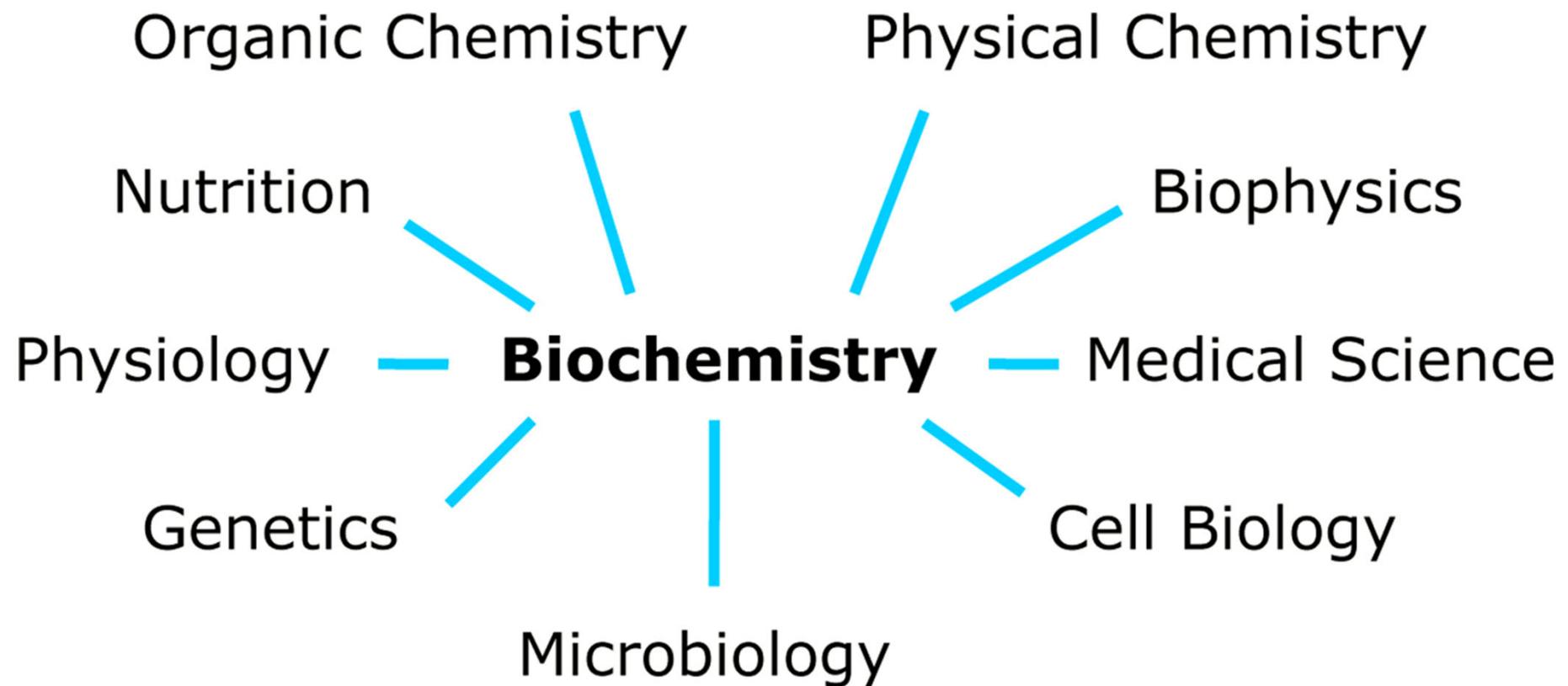
The Science of Biochemistry

- **Fermentation:** an ancient application of biochemistry: ~5000 years ago



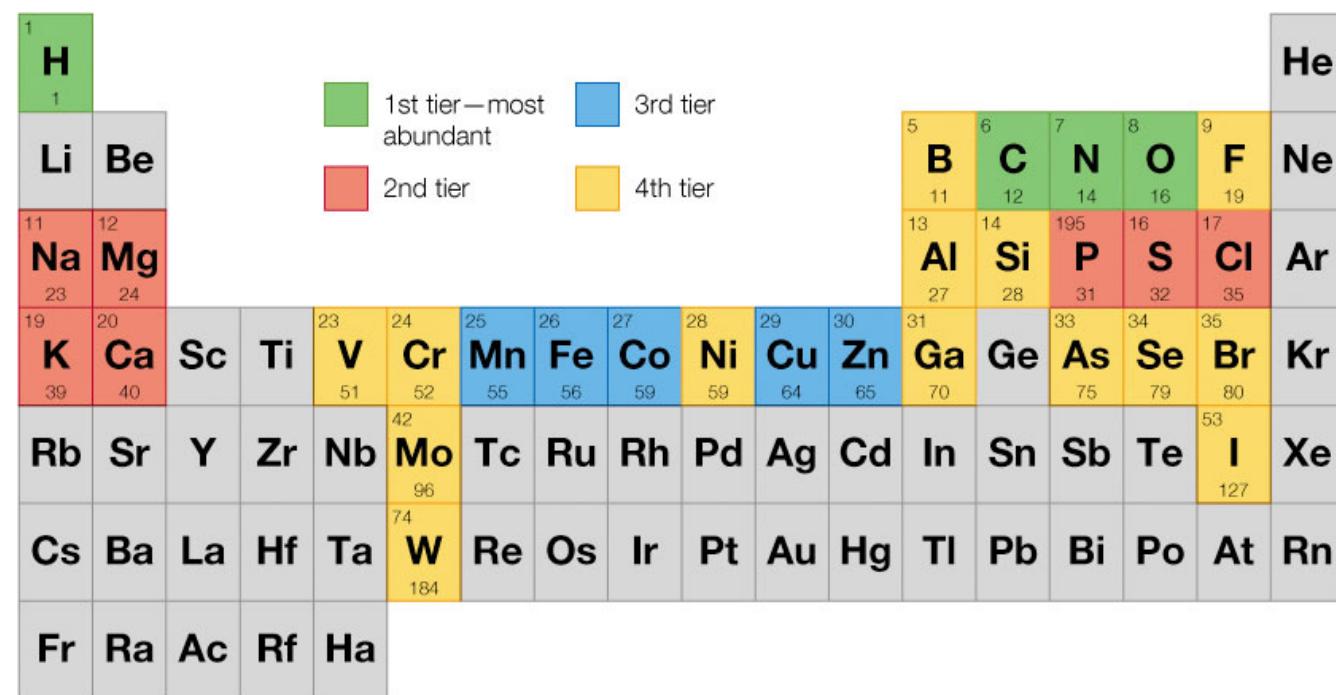
- 1897: Chemists Eduard and Hans Buchner found that **extracts from broken yeast cells** could carry out the entire process of fermentation of sugar into ethanol
- Biological catalysts, known as enzymes, promote biochemical reactions in living systems
- Synthesis of enzymes is controlled by genes
- Genes are the units of hereditary information, encoded in the structure of deoxyribonucleic acid (DNA)
- In 1953, James Watson and Francis Crick described the double-helical structure of DNA

Defining Biochemistry



The elements and molecules of living systems

- Living systems are primarily composed of **carbon (C)**, **hydrogen (H)**, **oxygen (O)**, and **nitrogen (N)**
 - These are also the most abundant elements
- However, other elements including **sulfur (S)** and **phosphorus (P)** and certain ions, such as **Na⁺**, **K⁺**, **Mg²⁺**, **Ca²⁺**, and **Cl⁻**, are essential for the existence of living systems on Earth as well
- Trace amounts** of many other elements are also essential



Origin of Living Systems

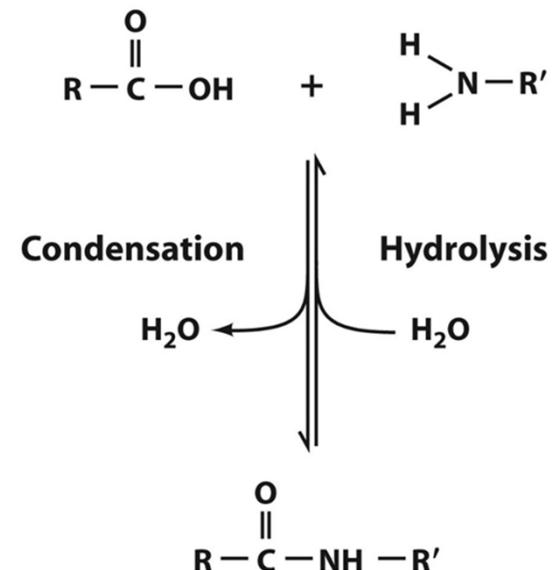
- Stanley Miller's 1953 experiments showed that certain building blocks of complex biomolecules, such as amino acids, could have been produced under early (abiotic) Earth conditions ("Primordial Soup")
- How cells, the basic unit of living systems, may have formed is still unknown
- However, many biochemists believe that ancient cells had already a ribonucleic acid (RNA)-based self-replication mechanism ("RNA World"): here RNA served as both genetic code (currently DNA) as well as enzymes (which are currently carried out by proteins)

Biological macromolecules

- Four major classes are essential for living systems based on cells:
 - A. Nucleic acids (DNA and RNA)
 - B. Proteins
 - C. Polysaccharides (polymeric carbohydrates)
 - D. Lipids
- Polymers of smaller organic molecule subunits
- Formed by condensation reactions (reverse reaction is hydrolysis): e.g. carboxylic acid and amine, forming an amide bond

| Macromolecule | Monomer | Linkage |
|----------------|----------------|-------------------|
| Nucleic acids | Nucleotide | Phosphodiester |
| Protein | Amino acid | Peptide (amide) |
| Polysaccharide | Monosaccharide | Glycoside (ether) |
| Lipids* | Fatty acids | Ester |

*Lipids are not polymeric, but form large assemblies



Distinguishing Characteristics of Living Systems

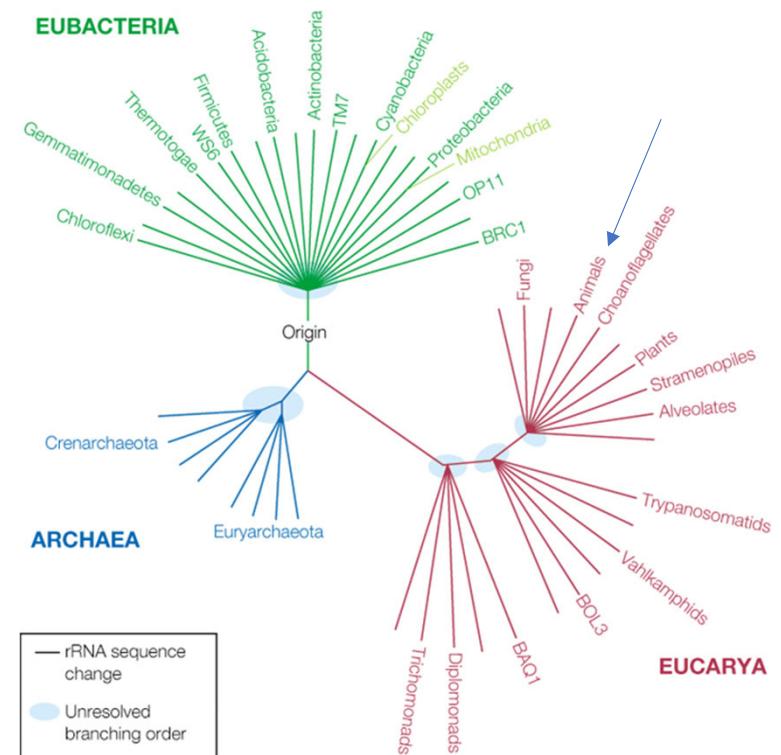
In 2002, Koshland summarized the essential attributes that distinguish living from nonliving things*:

- 1) **A program** (organized plan for constitution and regeneration; DNA)
- 2) **Improvisation** (changing the program as surroundings change; evolution)
- 3) **Compartmentation** (ability to be separate from environment; membranes)
- 4) **Energy** (ability to maintain order despite overall positive entropy)
- 5) **Regeneration** (compensation for environmental wear and tear; repair)
- 6) **Adaptability** (ability to respond to environmental changes)
- 7) **Seclusion** (operation of processes and pathways in isolation)

*From Koshland, D.E. (2002) **The seven pillars of life**. *Science* 295:2215-2216.

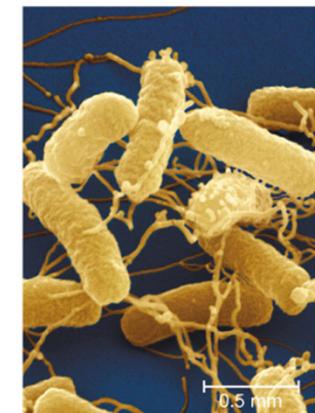
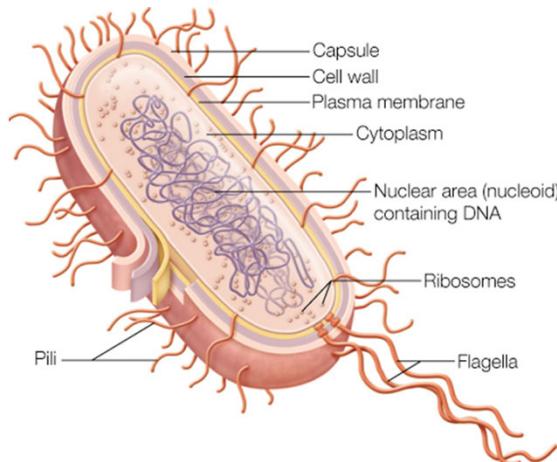
The Unit of Biological Organization: The Cell

- Cells are the universal unit of life
 - bacterial
 - archaeal
 - eukaryotic
- Although distinct in many aspects, bacterial and archaeal cells are summarized as prokaryotic cells
- Cell type correlates with phylogeny (“Molecular Tree of Life”)
 - Based on ribosomal RNA sequence comparisons



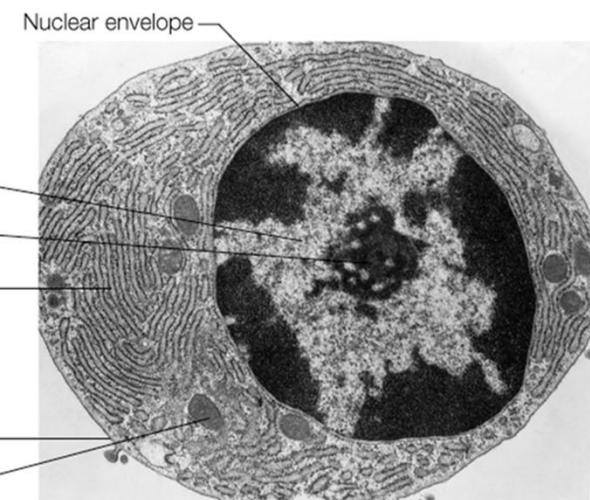
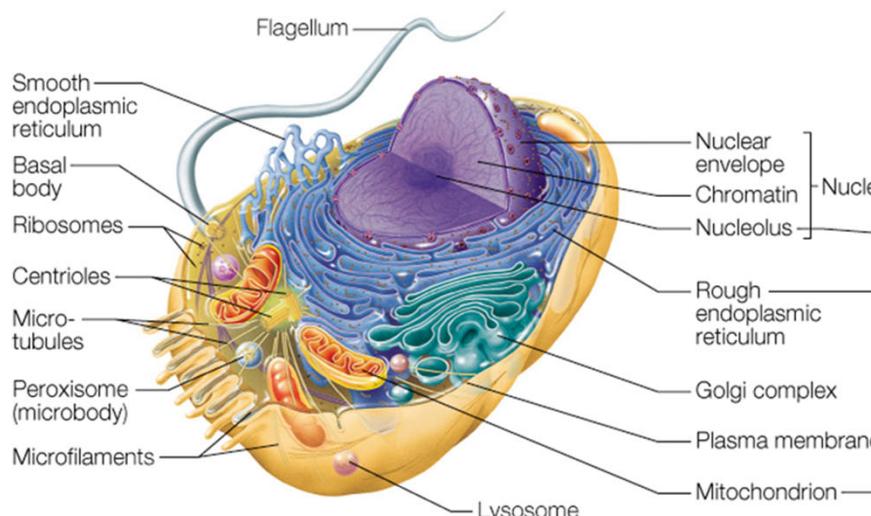
Types of cells

- **Prokaryotic cell**
- **Eukaryotic cells** have membrane-surrounded structures (organelles) within the cellular boundary



(a) Schematic view of a representative bacterial cell. The DNA molecule that constitutes most of the genetic material is coiled up in a region called the nucleoid, which shares the fluid interior of the cell (the cytoplasm) with ribosomes (which synthesize proteins), other particles, and a large variety of dissolved molecules. The cell is bounded by a plasma membrane, outside of which is usually a fairly rigid cell wall. Many bacteria also have a gelatinous outer capsule. Projecting from the surface may be pili, which attach the cell to other cells or surfaces, and one or more flagellae, which enable the cell to swim through a liquid environment.

(b) Scanning electron micrograph of *Salmonella*: rod-shaped Gram-negative enterobacteria that causes typhoid fever, paratyphoid fever, and food-borne illness.



(a) Typical animal cell. The accompanying photograph is an electron micrograph of a representative animal cell, a white blood cell.

New Research Methods in Biochemistry

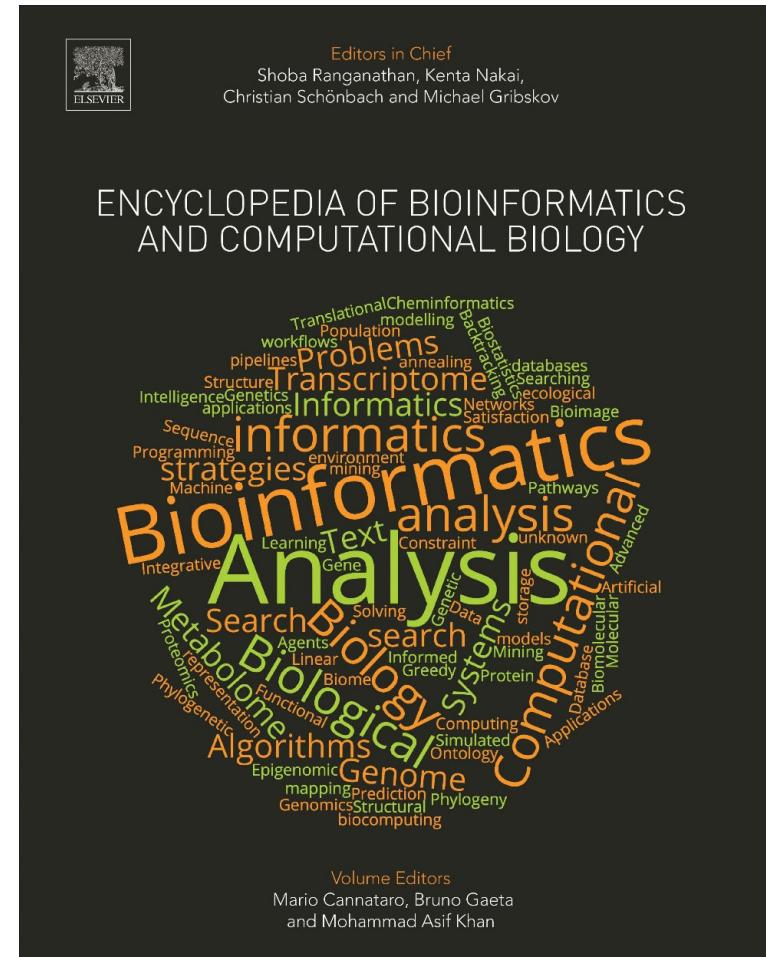
| | | | | | |
|------|--|------|--|------|---|
| 2015 | • Cryo-electron microscopy • CRISPR-Cas9 technology | 1985 | • Pulsed field electrophoresis • Transgenic animals • Amplification of DNA: polymerase chain reaction • Automated oligonucleotide synthesis • Site-directed mutagenesis of cloned genes • Automated micro-scale protein sequencing • Rapid DNA sequence determination • Monoclonal antibodies • Southern blotting • Two-dimensional gel electrophoresis • Gene cloning | 1965 | • High-performance liquid chromatography • Polyacrylamide gel electrophoresis • Solution hybridization of nucleic acids • X-ray crystallographic protein structure determination • Zone sedimentation velocity centrifugation • Equilibrium gradient centrifugation • Liquid scintillation counting |
| 2010 | • Synthetic biology • RNA-sequence analysis • Chromatin immunoprecipitation/sequencing • Induced pluripotent cells • Second generation DNA sequence analysis | 1980 | • Automated oligonucleotide synthesis • Site-directed mutagenesis of cloned genes • Automated micro-scale protein sequencing • Rapid DNA sequence determination • Monoclonal antibodies • Southern blotting • Two-dimensional gel electrophoresis • Gene cloning | 1960 | • First determination of the amino acid sequence of a protein • X-ray diffraction of DNA fibers |
| 2005 | • Proteomic analysis with mass spectrometry | 1975 | • Restriction cleavage mapping of DNA molecules • Rapid methods for enzyme kinetics | 1955 | • Radioisotopic tracers used to elucidate reactions |
| 2000 | • Genetic code expansion • Gene analysis on microchips • Single-molecule dynamics | 1970 | | 1950 | |
| 1995 | • Targeted gene disruption • In vivo NMR | | | 1945 | |
| 1990 | • Atomic force microscopy • Scanning tunneling microscopy | | | | |



Biochemistry and the Information Explosion

Bioinformatics

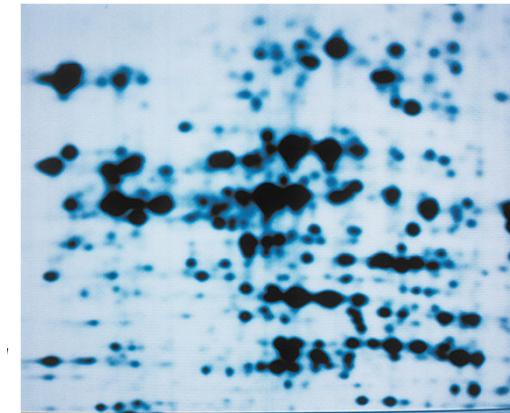
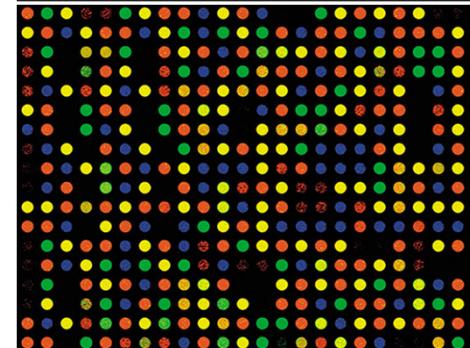
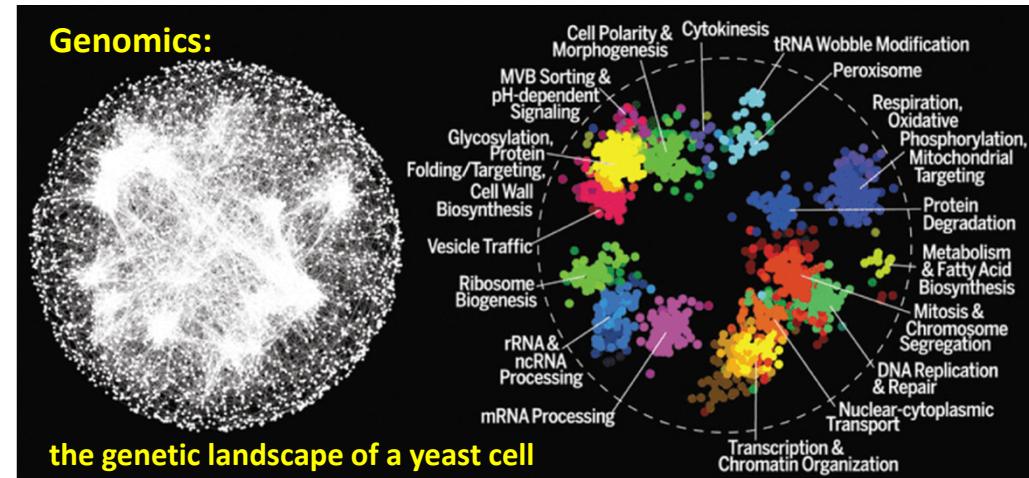
- New scientific tools and techniques can generate and analyze increasing amount of biochemical and molecular biological information
 - **Bioinformatics** can be considered as **information science applied to biology**. Examples are
 - 1) mathematical analysis of DNA sequence data
 - 2) *in silico* analysis of protein structure and function
 - 3) computer simulation of metabolic pathways
 - 4) analysis of potential drug targets (enzymes or receptors) for structure-based drug design



Biochemistry and the Information Explosion - 2

Omics

- **Genetics** concerns itself with the location, expression, and function of individual genes or small groups of genes
- **Genomics** expands the genetics approach and concerns itself with the entire genome, the totality of genetic information in an organism
 - Some of the broader goals of genomics are to
 - 1) determine the nucleotide sequence of the whole genome (of an organism)
 - 2) assess the expression and function of each gene
 - 3) understand the evolutionary relationships among genes in the same genome and with genomes of different organisms
- Examples of other established “Omics” are **Proteomics, Transcriptomics, Metabolomics, and Interactomics**



Transcriptomics: the abundance of gene-specific mRNAs indicated by microarrays

Proteomics: separation of a large number of proteins from a cellular extract by two-dimensional gel electrophoresis

Biochemistry - Summary

- The aim of biochemistry is to understand and explain living systems in molecular terms
- Biochemistry bridges biological and chemical sciences on the level of molecules present in living systems
- Living systems are composed of cells, which can be divided into three major types: bacterial, archaeal, and eukaryotic
- Biochemistry is an experimental science and uses a variety of different tools and techniques, some of which generate large amount of information

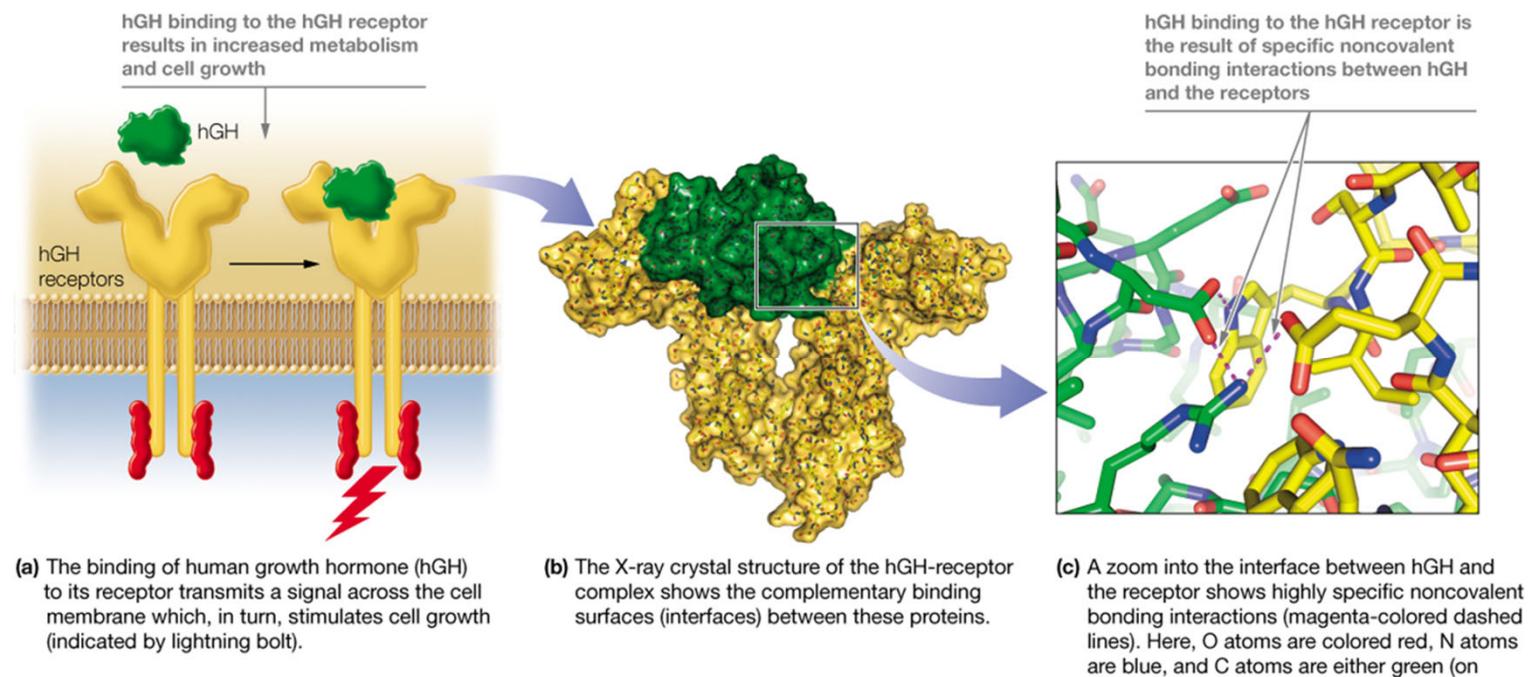
The Chemical Foundation of Life: Weak Interactions in an Aqueous Environment

AAM: Chapter 2

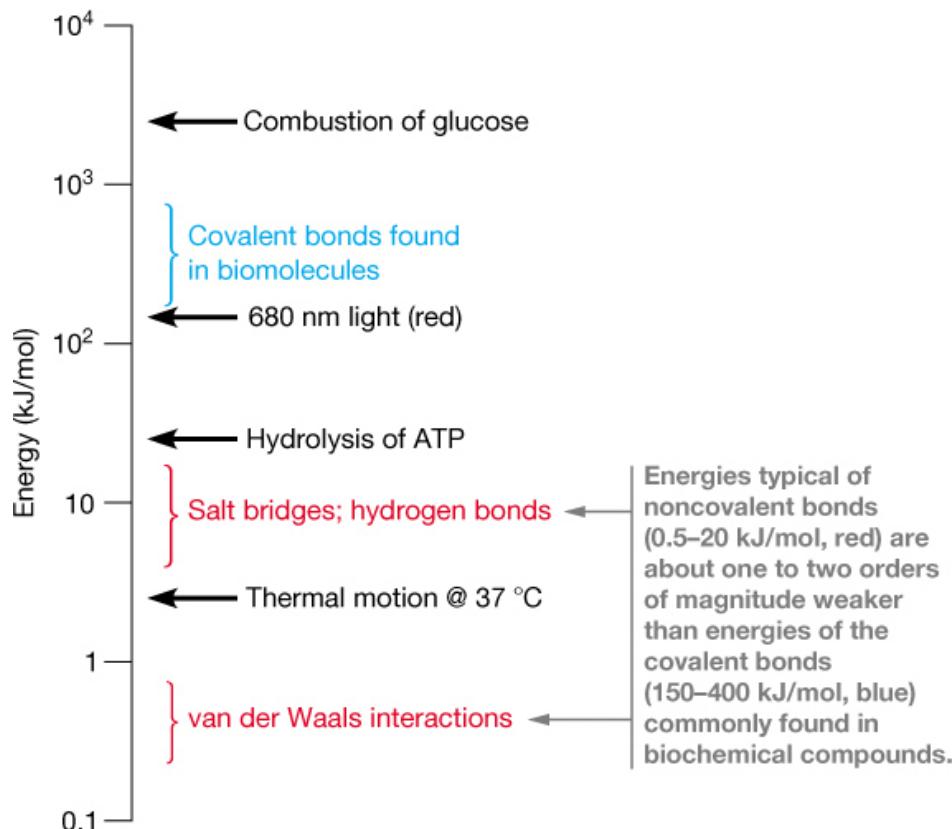
- The Importance of Noncovalent Interactions in Biochemistry
- The Nature of Noncovalent Interactions
- The Role of Water in Biological Processes

Noncovalent Interactions Define the Structure and Function of Biomolecules

- a) Binding of hormone to its receptor
- b) The binding pocket shows surfaces that complement each other.
- c) The binding occurs due to formation of non-covalent bonds.



Noncovalent Bonds vs. Covalent Bonds



- Covalent bonds are strong and not easily broken
- Noncovalent bonds are weak, and thus can be continually broken and reformed

TABLE 2.1 Energies of some noncovalent interactions in biomolecules

| Type of Interaction | Approximate Energy (kJ/mol) |
|---------------------|-----------------------------|
| Charge–charge | 13 to 17 |
| Hydrogen bond | 2 to 21 |
| van der Waals | 0.4 to 0.8 |

Source: Data from S. K. Burley and G. A. Petsko, Weakly polar interactions in proteins, *Advances in Protein Chemistry* (1988) 39:125–189.

The Nature of Noncovalent Interactions

All noncovalent interactions are electrostatic in nature:

(a) Charge-charge interactions between two charged particles – simplest

(g) Hydrogen bonding is very important for the structure and properties of biomolecules

- between a hydrogen atom covalently bonded to another atom (e.g., O-H or N-H) and a pair of nonbonded electrons on a separate (e.g., O or N) atom

| Type of Interaction | Model | Example | Dependence of Energy on Distance |
|--------------------------------|-------|---|----------------------------------|
| (a) Charge-charge | | --NH_3^+ $\text{O}=\text{C}\text{--}$ | $1/r$ |
| (b) Charge-dipole | | --NH_3^+ $\text{H}\text{--O}^{\delta-}\text{O}^{\delta+}\text{H}$ | $1/r^2$ |
| (c) Dipole-dipole | | $\text{H}\text{--O}^{\delta-}\text{H}$ $\text{H}\text{--O}^{\delta-}\text{H}$ | $1/r^3$ |
| (d) Charge-induced dipole | | --NH_3^+ $\text{C}_6\text{H}_5\text{--}$ | $1/r^4$ |
| (e) Dipole-induced dipole | | $\text{H}\text{--O}^{\delta-}\text{H}$ $\text{C}_6\text{H}_5\text{--}$ | $1/r^5$ |
| (f) Dispersion (van der Waals) | | | $1/r^6$ |
| (g) Hydrogen bond | | $\text{N}-\text{H}\cdots\text{O}=\text{C}\text{--}$ | Bond length is fixed |



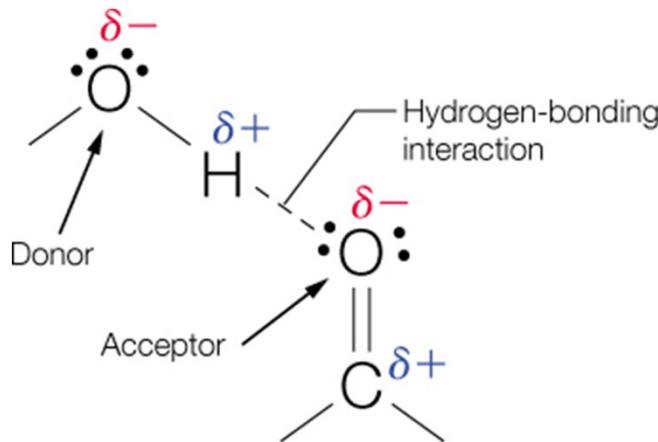
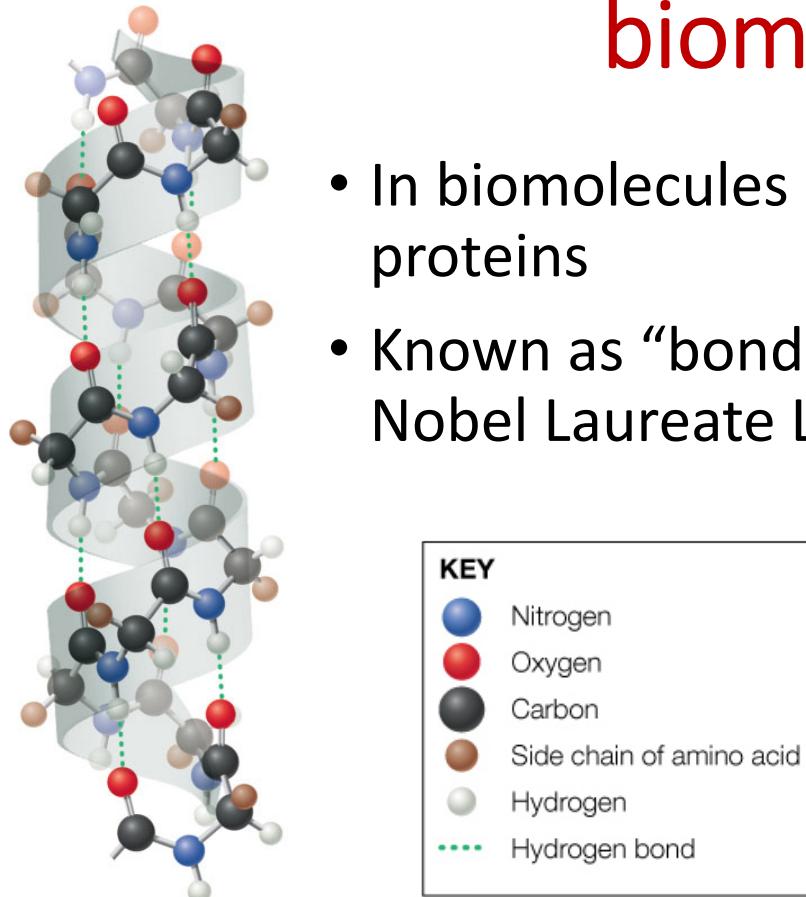


TABLE 2.3 Major types of hydrogen bonds found in biomolecular interactions

| Donor . . . Acceptor | Distance between Donor and Acceptor (\AA) | Comment |
|---|--|--|
| $-\text{O}-\text{H} \cdots \text{O}-\text{H}$ | 2.8 ± 0.1 | H bond formed in water |
| $-\text{O}-\text{H} \cdots \text{O}=\text{C}$ | 2.8 ± 0.1 | Bonding of water to other molecules often involves these |
| $-\text{N}-\text{H} \cdots \text{O}-\text{H}$ | 2.9 ± 0.1 | |
| $-\text{N}-\text{H} \cdots \text{O}=\text{C}$ | 2.9 ± 0.1 | Very important in protein and nucleic acid structures |
| $-\text{N}-\text{H} \cdots \text{N}$ | 3.1 ± 0.2 | |
| $-\text{N}-\text{H} \cdots \text{S}$ | 3.7 | Relatively rare; weaker than above |

Hydrogen bonding in biomolecules

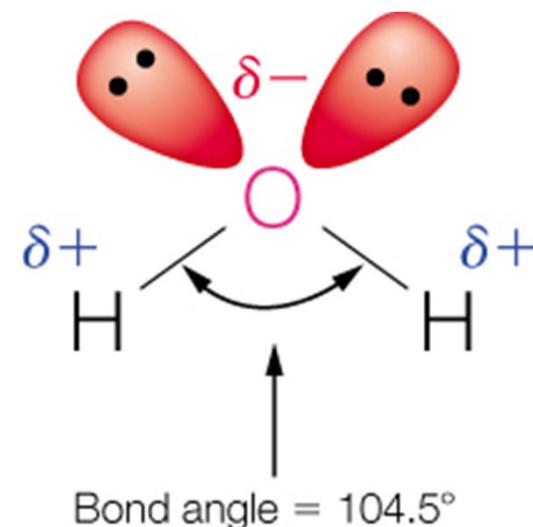
- In biomolecules such as proteins
- Known as “bonds of life” – Nobel Laureate Linus Pauling



The Role of Water in Biological Processes

Unique properties of water that make it suitable as the medium of life

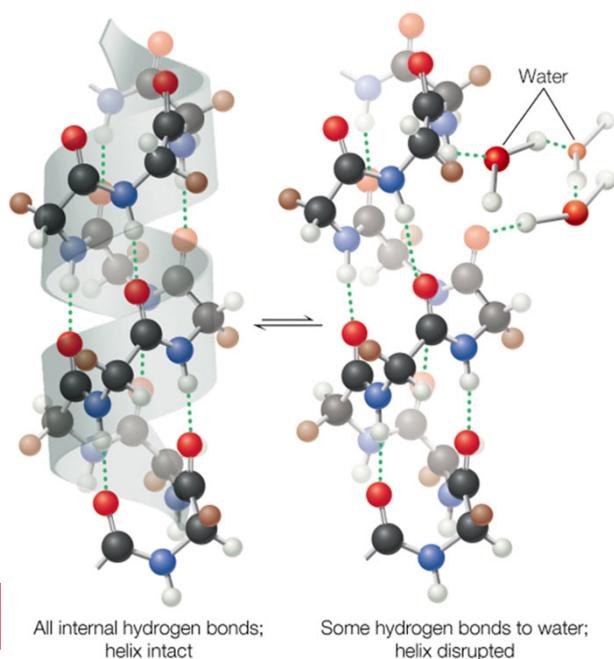
- a permanent dipole
 - two H bond donor sites and two H bond acceptor sites
- high heat capacity
- liquid with high density
- relatively high dielectric constant, making it capable of solvating ionic compounds



Solvation of hydrophilic, hydrophobic and amphipathic substances

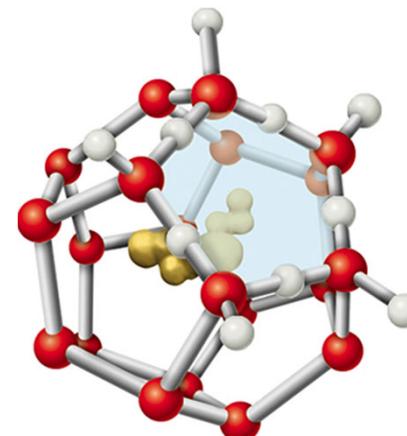
Hydrophilic:

- Solvent can compete with intramolecular H-bond donors/acceptors



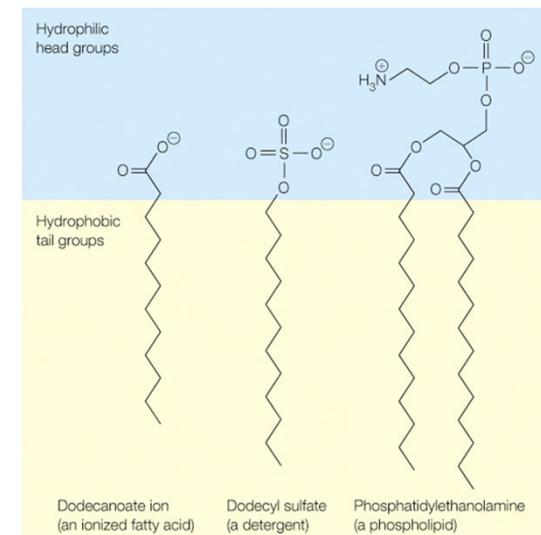
Hydrophobic:

- Solvent can form “clathrate”(cage) structures to surround nonpolar surface areas
- “Hydrophobic effect”—stabilizes protein structure by driving apolar groups together



Amphipathic:

- An amphipathic substance has hydrophilic and hydrophobic parts and can form a monolayer, a micelle, or a bilayer.
- Bilayers are typical of biological membranes



Take home message

- Water is essential for all forms of life.
- Every organism is 70-90% water!
- Normal human metabolic activity requires min. 65% water!
- Water is an excellent solvent and is therefore the medium of the majority of biochemical reactions.
- Water causes ionization of polar molecules, critical for the function of:
 - amino acids and proteins
 - nucleotides and nucleic acids
 - even phospholipids and membranes
- Water solvates hydrophilic, hydrophobic and amphipathic molecules.

Amino Acids 1: Intro and Acid-Base Properties

Shoba Ranganathan

Applied Biosciences

T: 02 9850 6262; E: shoba.ranganathan@mq.edu.au

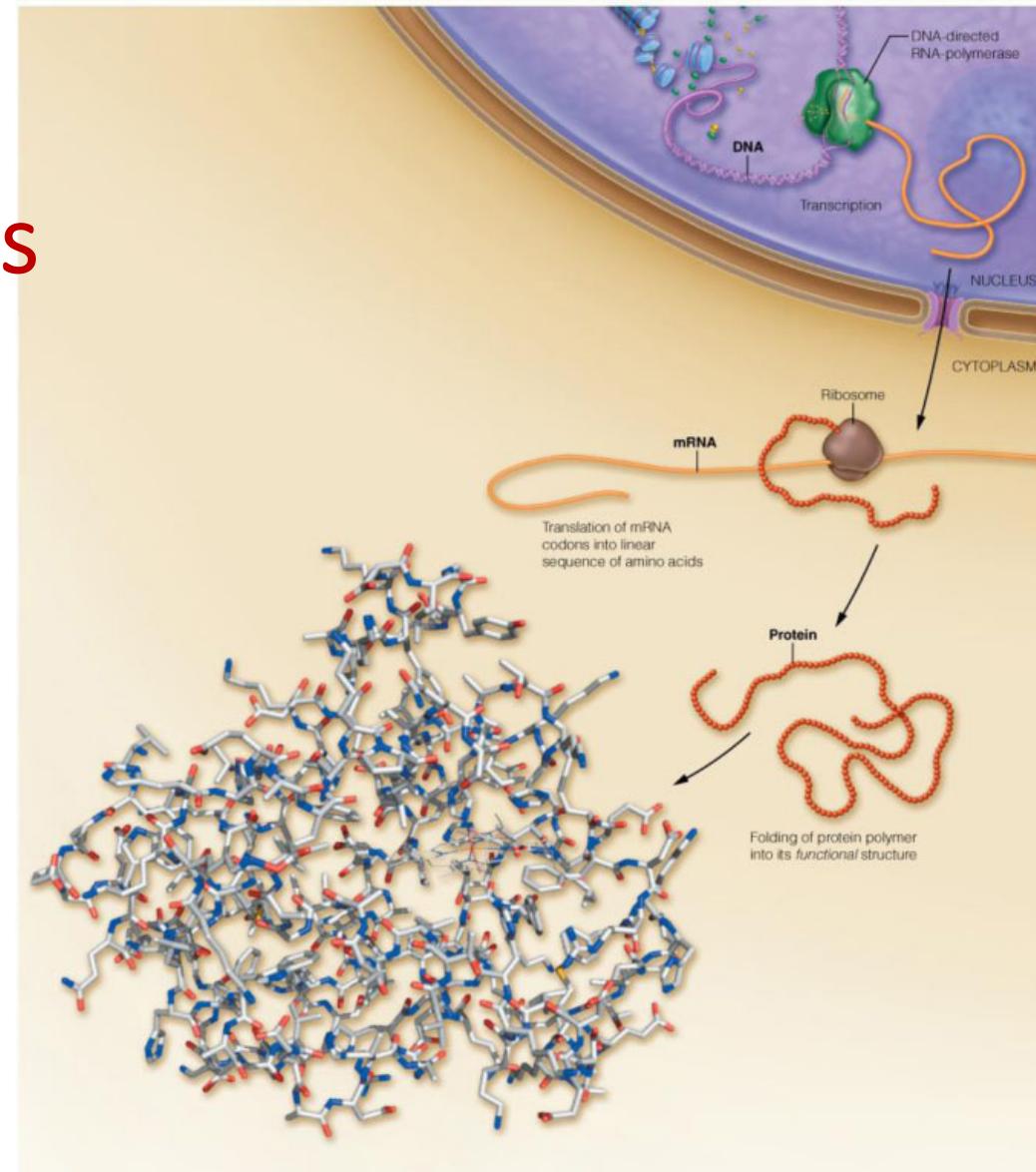
Objectives

- What are amino acids?
 - How are amino acids linked to form polymeric chains (peptides or proteins)?
- What are the acid-base properties of specific amino acids as well as the peptides and proteins they form?
 - Theory for Prac 1 starting Week 2.

AAM: Chapters 4 and 2

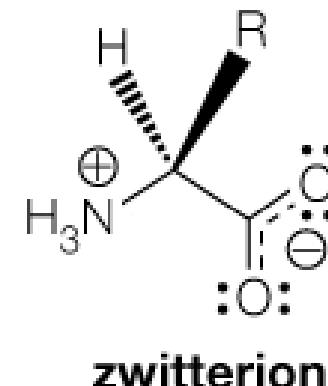
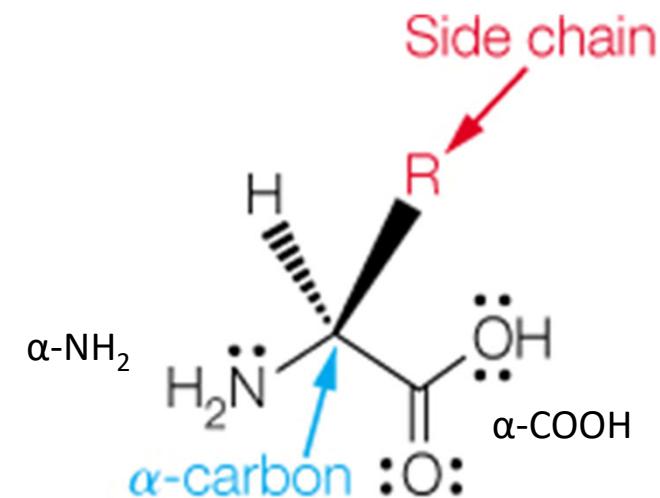
Amino acids are the building blocks of proteins

- DNA is transcribed to form messenger RNA (mRNA) in the nucleus
- mRNA is exported to the cytoplasm, where it is bound to the ribosome
- mRNA is translated into **a linear sequence of amino acids** that folds into a 3D structure



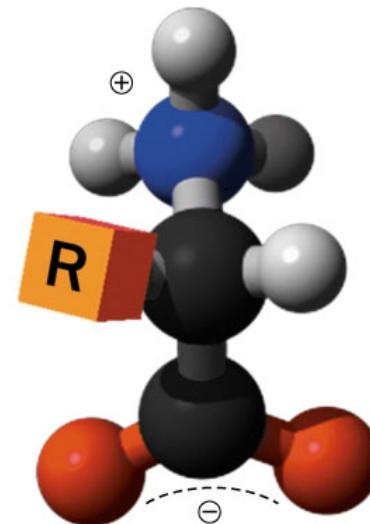
What are amino acids?

- The **α -Carbon** is central to an amino acid: amino group, carboxylic acid, and amino acid-specific side chain (R group) are attached to it
 - when R is not a hydrogen atom, the α -carbon is an **asymmetric** (chiral) centre
- At neutral pH (~7), the amino group nitrogen is protonated and the carboxylic acid is deprotonated to yield a “zwitterion” which has 0 net charge!

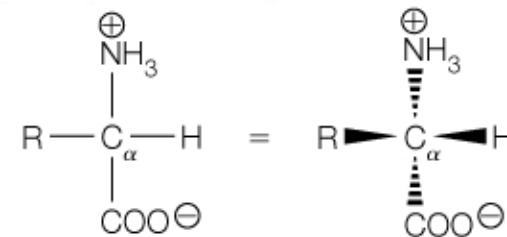


α -Amino Acid Stereochemistry - 1

- The ball and stick model shows the 3D arrangement of the atoms in an α -amino acid
- When four different groups are attached to a carbon atom, that atom is said to be chiral, or a stereocenter, or an asymmetric carbon
- The Fischer projection in the bottom figure (right) is a 2D stereochemistry representation



(a) This ball-and-stick model shows the three-dimensional arrangement of the atoms, with C atoms in black, H atoms in gray, the N atom in blue, and O atoms in red. The α -carbon is asymmetric, with tetrahedral bonding geometry. The variable R group is shown as an orange cube.



(b) In a Fischer projection (left), the horizontal bonds project toward the viewer, and the vertical bonds project away from the viewer. This orientation of bonds in the Fischer projection is represented on the right by solid and dashed wedges, respectively.

FIGURE 5.4 Three-dimensional representations of α -amino acids.

Module 1: Building Blocks of Biochemistry

α -Amino Acid Stereochemistry - 2

- All α -amino acids (20), except glycine, contain an asymmetric α -carbon
- When R = -CH₃ (methyl group), we have the amino acid alanine
- L-alanine is the mirror image of D-alanine; they are enantiomers

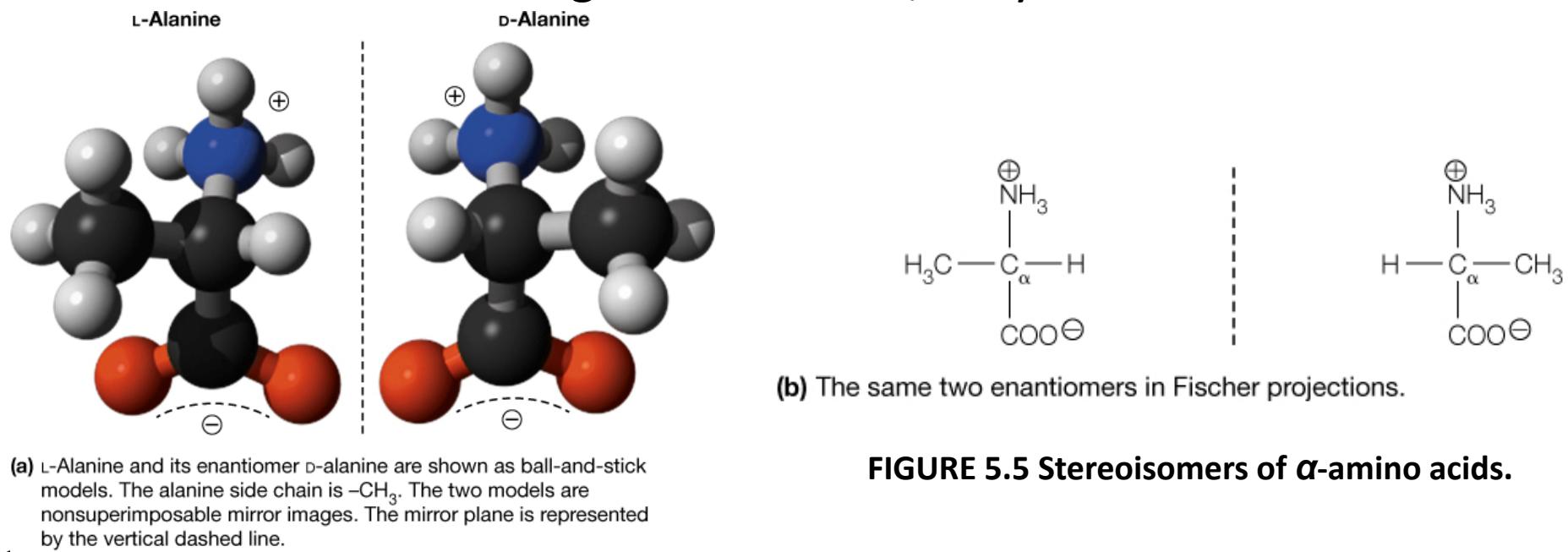
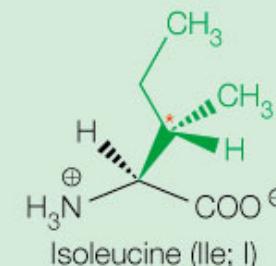
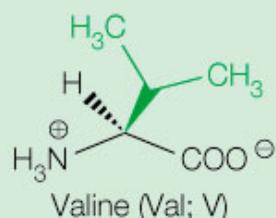
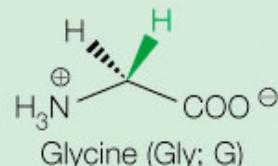


FIGURE 5.5 Stereoisomers of α -amino acids.

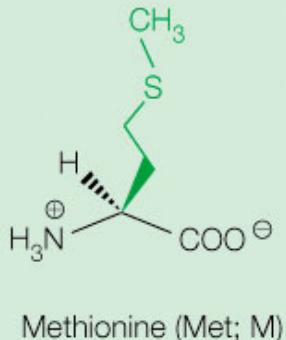
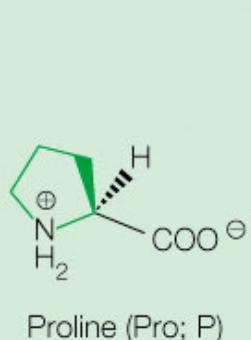
Classification of the 20 Naturally Occurring Amino Acids - 1

FIGURE 5.3 The 20 common amino acids found in proteins.

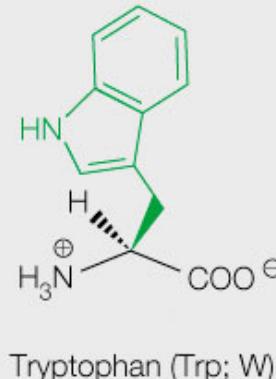
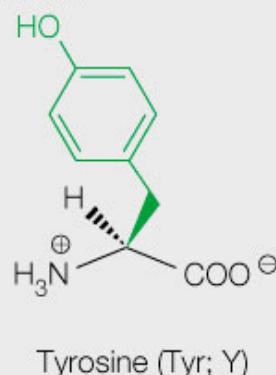
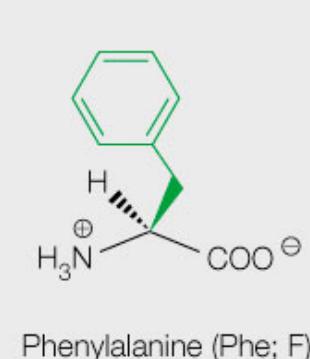
Nonpolar Aliphatic Amino Acids



Nonpolar Amino Acids

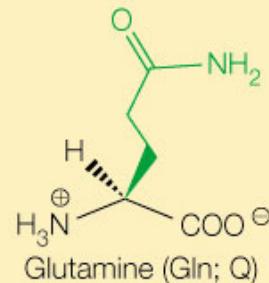
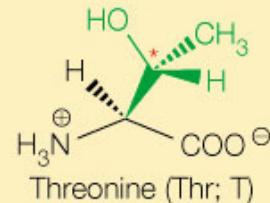
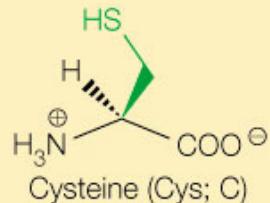
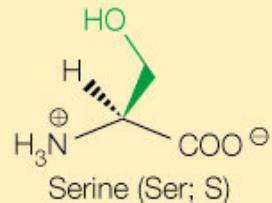


Nonpolar Aromatic Amino Acids

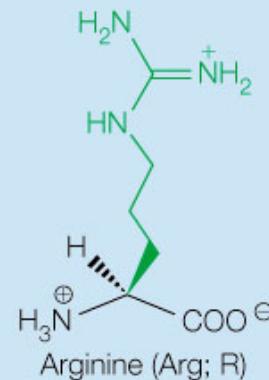
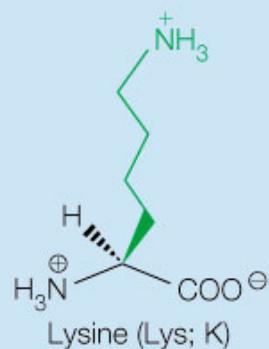
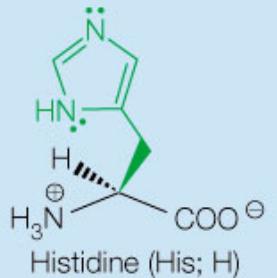


Classification of the 20 Naturally Occurring Amino Acids - 2

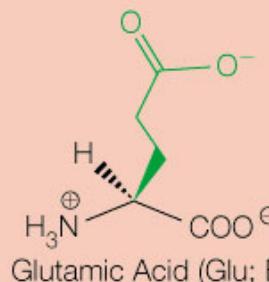
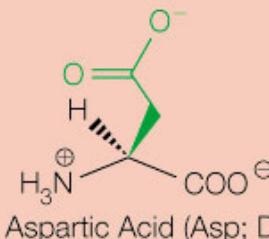
Polar Amino Acids



Positively Charged Polar Amino Acids



Negatively Charged Polar Amino Acids



General Properties of Amino Acids

TABLE 5.1 Properties of the common amino acids found in proteins

| Name | Abbreviations: 1- and 3-letter codes | pK _a of α -COOH Group ^a | pK _a of α -NH ₃ ⁺ Group ^a | pK _a of Ionizing Side Chain ^a | Residue ^b Mass (daltons) | Occurrence ^c in Proteins (mol %) |
|---------------|--------------------------------------|--|--|---|-------------------------------------|---|
| Alanine | A, Ala | 2.3 | 9.7 | — | 71.08 | 8.7 |
| Arginine | R, Arg | 2.2 | 9.0 | 12.5 | 156.20 | 5.0 |
| Asparagine | N, Asn | 2.0 | 8.8 | — | 114.11 | 4.2 |
| Aspartic acid | D, Asp | 2.1 | 9.8 | 3.9 | 115.09 | 5.9 |
| Cysteine | C, Cys | 1.8 | 10.8 | 8.3 | 103.14 | 1.3 |
| Glutamine | Q, Gln | 2.2 | 9.1 | — | 128.14 | 3.7 |
| Glutamic acid | E, Glu | 2.2 | 9.7 | 4.2 | 129.12 | 6.6 |
| Glycine | G, Gly | 2.3 | 9.6 | — | 57.06 | 7.9 |
| Histidine | H, His | 1.8 | 9.2 | 6.0 | 137.15 | 2.4 |
| Isoleucine | I, Ile | 2.4 | 9.7 | — | 113.17 | 5.5 |
| Leucine | L, Leu | 2.4 | 9.6 | — | 113.17 | 8.9 |
| Lysine | K, Lys | 2.2 | 9.0 | 10.0 | 128.18 | 5.5 |
| Methionine | M, Met | 2.3 | 9.2 | — | 131.21 | 2.0 |
| Phenylalanine | F, Phe | 1.8 | 9.1 | — | 147.18 | 4.0 |
| Proline | P, Pro | 2.0 | 10.6 | — | 97.12 | 4.7 |
| Serine | S, Ser | 2.2 | 9.2 | — | 87.08 | 5.8 |
| Threonine | T, Thr | 2.6 | 10.4 | — | 101.11 | 5.6 |
| Tryptophan | W, Trp | 2.4 | 9.4 | — | 186.21 | 1.5 |
| Tyrosine | Y, Tyr | 2.2 | 9.1 | 10.1 | 163.18 | 3.5 |
| Valine | V, Val | 2.3 | 9.6 | — | 99.14 | 7.2 |

^aApproximate values found for side chains on the free amino acids. W. P. Jencks and J. Regenstein (1976) Ionization constants of acids and bases in *Handbook of Biochemistry and Molecular Biology*, 3rd ed., G. Fasman (ed.), CRC Press, Boca Raton, FL.

^bTo obtain the mass of the amino acid itself, add the mass of a molecule of water, 18.02 daltons. The values given are for neutral side chains; slightly different masses are observed at pH values where protons have been gained or lost from the side chains.

^cAverage for a large number of proteins. Individual proteins can show large deviations from these values. Data from J. M. Otaki, M. Tsutsumi, T. Gotoh, and H. Yamamoto, Secondary structure characterization based on amino acid composition and availability in proteins (2010) *Journal of Chemical Information and Modeling* 50:690–700 © 2010 American Chemical Society.



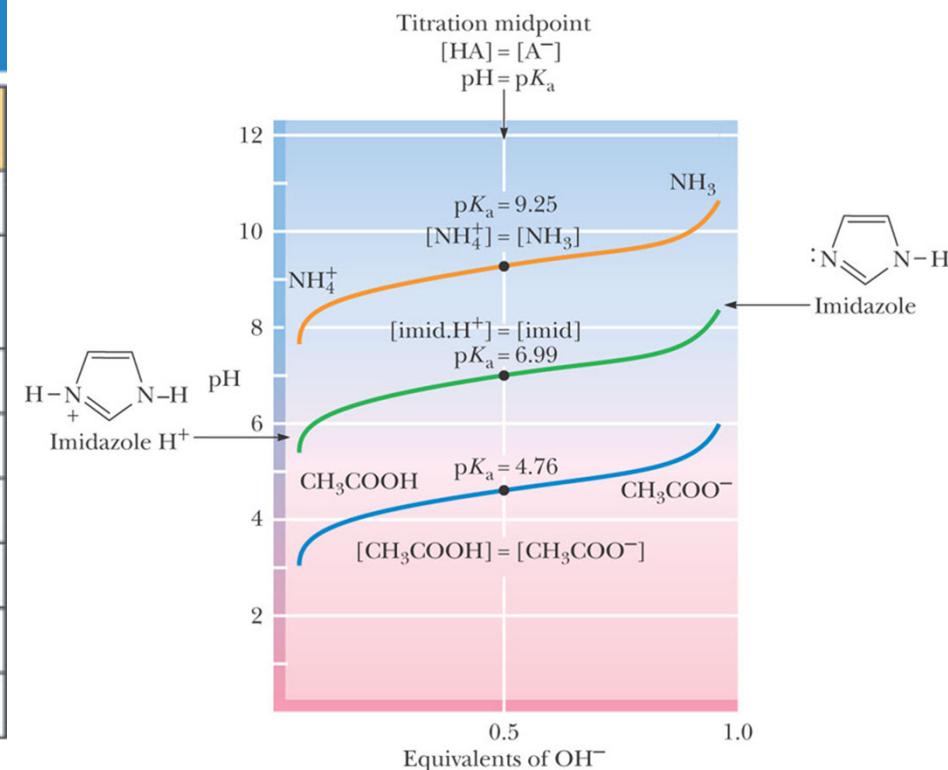
Ionizable groups of Amino Acids

pK_a values of ionizable amino acid groups: Prac 1 theory!

TABLE 5.2 Typical ranges observed for pK_a values of ionizable groups in proteins

| Group Type | Typical pK_a Range ^a |
|--|-----------------------------------|
| α -Carboxyl | 3.5–4.0 |
| Side-chain carboxyl (aspartic and glutamic acids) | 4.0–4.8 |
| Imidazole (histidine) | 6.5–7.4 |
| Cysteine (–SH) | 8.5–9.0 |
| Phenolic (tyrosine) | 9.5–10.5 |
| α -Amino | 8.0–9.0 |
| Side-chain amino (lysine) | 9.8–10.4 |
| Guanidinyl (arginine) | ~12 |

^aValues outside these ranges are observed. For example, side-chain carboxyls have been reported with pK_a values as high as 7.3.



Amino acids - summary

- The 20 standard amino acids share a common structure but differ in their side chains.
- Some amino acid side chains contain ionizable groups whose pK values may vary.

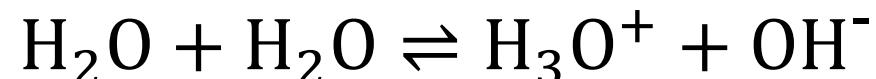
Acid-Base Equilibria

Acid and Bases: Proton Donors and Acceptors

- The behavior of molecules in biochemical processes occurring in aqueous environment depends on their state of ionization, particularly protonation
- **Acids are proton (H^+) donors and bases are proton acceptors** (“Brønsted-Lowry definition”)
- A strong acid dissociates (almost) completely into a proton and a conjugate base, while a weak acid dissociates only partially
 - the dissociated proton is transferred to water (H_2O) to yield a hydronium ion (H_3O^+)

Acid and Bases: Proton Donors and Acceptors

- A strong base ionizes entirely, generating an OH^- (hydroxide ion) and a conjugate acid, while a weak base ionizes only partially
- **Water** itself can act as both a weak acid or weak base
- It is **amphiprotic** and **autoionizes**:



Water as a Proton Donor and Acceptor

- At equilibrium:

$$K_w = \frac{(a_{H^+})(a_{\text{-OH}})}{(a_{H_2O})} = 10^{-14}$$

$$K_w \simeq \frac{[H^+][\text{-OH}]}{1} = 10^{-14}$$

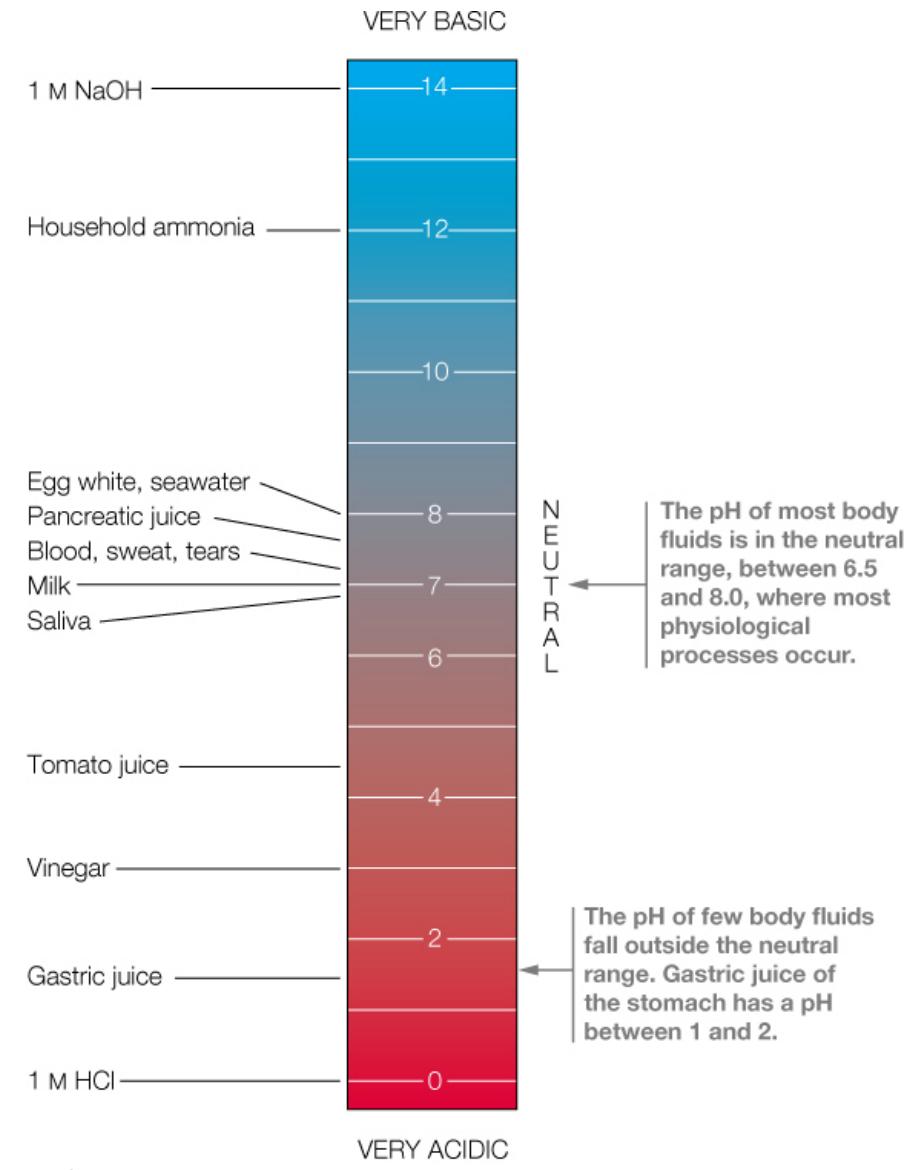
- K_w is the ion product (or ion-product constant) for water

The pH Scale and the Physiological pH Range

- The hydrogen ion (H^+ , or hydronium ion, H_3O^+) concentration can be expressed in terms of pH:

$$pH = -\log(a_{H^+}) \cong -\log[H^+]$$

- The higher** $[H^+]$, the lower the pH (and vice versa)
 - Acidic solution have a low pH (<7), while basic solutions have a high pH (>7)
- Most biological reactions take place between **pH 6.5 and 8.0**, which is the physiological pH range



Molecular charge depends on pH

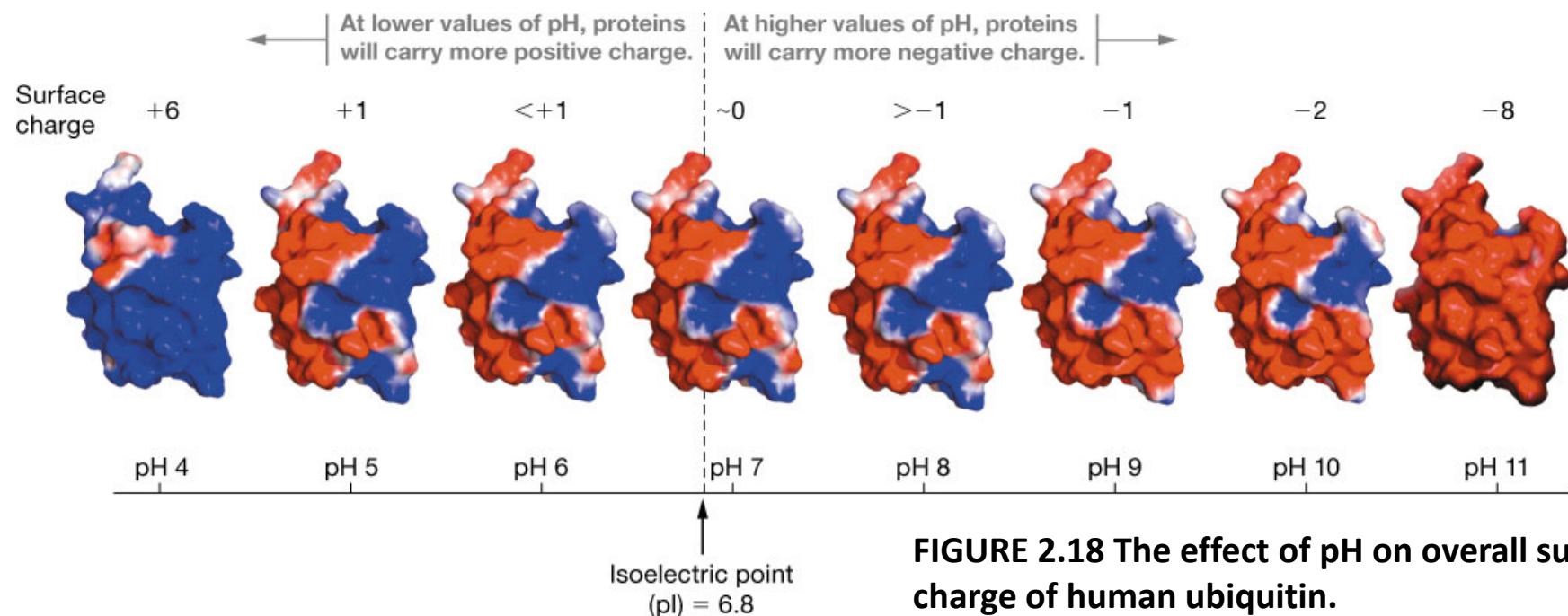


FIGURE 2.18 The effect of pH on overall surface charge of human ubiquitin.

The relationship between solution pH and molecular charge explains many features of biomolecular interactions

Weak Acid and Base Equilibria: K_a and pK_a

- Consider the dissociation of a weak acid:



- The equilibrium for the dissociation of a weak acid (K_a or the acid dissociation constant) is defined as:

$$K_a = \frac{[\text{H}^+][\text{A}^-]}{[\text{HA}]}$$

- The strength of acids is expressed as $\text{p}K_a = -\log K_a$

K_a and pK_a for some common weak acids

TABLE 2.6 Some weak acids and their conjugate bases

| Acid (Proton Donor) | | Conjugate Base (Proton Acceptor) | | pK_a | K_a (M) |
|---|----------------------|--|-----------------|--------|------------------------|
| HCOOH Formic acid | \rightleftharpoons | HCOO ⁻ Formate ion | +H ⁺ | 3.75 | 1.78×10^{-4} |
| CH ₃ COOH Acetic acid | \rightleftharpoons | CH ₃ COO ⁻ Acetate ion | +H ⁺ | 4.76 | 1.74×10^{-5} |
| $\begin{array}{c} \text{OH} \\ \\ \text{CH}_3\text{CH}-\text{COOH} \\ \text{Lactic acid} \end{array}$ | \rightleftharpoons | $\begin{array}{c} \text{OH} \\ \\ \text{CH}_3\text{CH}-\text{COO}^- \\ \text{Lactate ion} \end{array}$ | +H ⁺ | 3.86 | 1.38×10^{-4} |
| H ₃ PO ₄ Phosphoric acid | \rightleftharpoons | H ₂ PO ₄ ⁻ Dihydrogen phosphate ion | +H ⁺ | 2.14 | 7.24×10^{-3} |
| H ₂ PO ₄ ⁻ Dihydrogen phosphate ion | \rightleftharpoons | HPO ₄ ²⁻ Monohydrogen phosphate ion | +H ⁺ | 6.86 | 1.38×10^{-7} |
| HPO ₄ ²⁻ Monohydrogen phosphate ion | \rightleftharpoons | PO ₄ ³⁻ Phosphate ion | +H ⁺ | 12.4 | 3.98×10^{-13} |
| H ₂ CO ₃ Carbonic acid | \rightleftharpoons | HCO ₃ ⁻ Bicarbonate ion | +H ⁺ | 6.3* | $5.1 \times 10^{-7*}$ |
| HCO ₃ ⁻ Bicarbonate ion | \rightleftharpoons | CO ₃ ²⁻ Carbonate ion | +H ⁺ | 10.25 | 5.62×10^{-11} |
| C ₆ H ₅ OH Phenol | \rightleftharpoons | C ₆ H ₅ O ⁻ Phenolate ion | +H ⁺ | 9.89 | 1.29×10^{-10} |
| NH ₄ ⁺ Ammonium ion | \rightleftharpoons | NH ₃ Ammonia | +H ⁺ | 9.25 | 5.62×10^{-10} |

*Apparent pK_a and K_a values (see text for explanation).



Acid–Base Titrations

$$K_a = \frac{[H^+][A^-]}{[HA]}$$

$$pH = -\log[H^+] \quad \text{and} \quad pK_a = -\log K_a$$

$$-\log[H^+] = -\log K_a + \log \frac{[A^-]}{[HA]} \quad \text{or} \quad pH = pK_a + \log \frac{[A^-]}{[HA]}$$

- The equation to the right is known as the **Henderson–Hasselbalch equation**

Acid–Base Titrations can help determine pK_a

- The Henderson–Hasselbalch equation relates pH, pK_a and $[A^-]/[HA]$
- When:
 - $pH < pK_a$, $[HA] > [A^-]$
 - $pH = pK_a$, $[HA] = [A^-]$
 - $pH > pK_a$, $[HA] < [A^-]$
- A titration of a weak acid can determine when $pH = pK_a$, $[HA] = [A^-]$

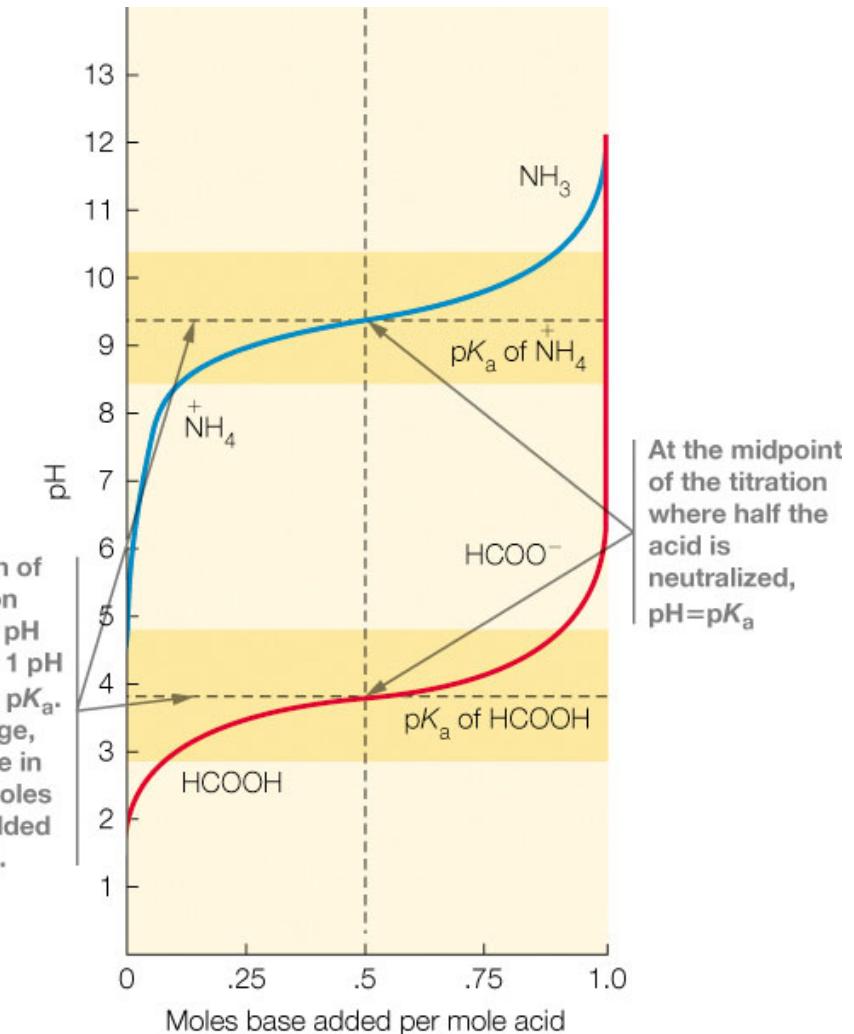
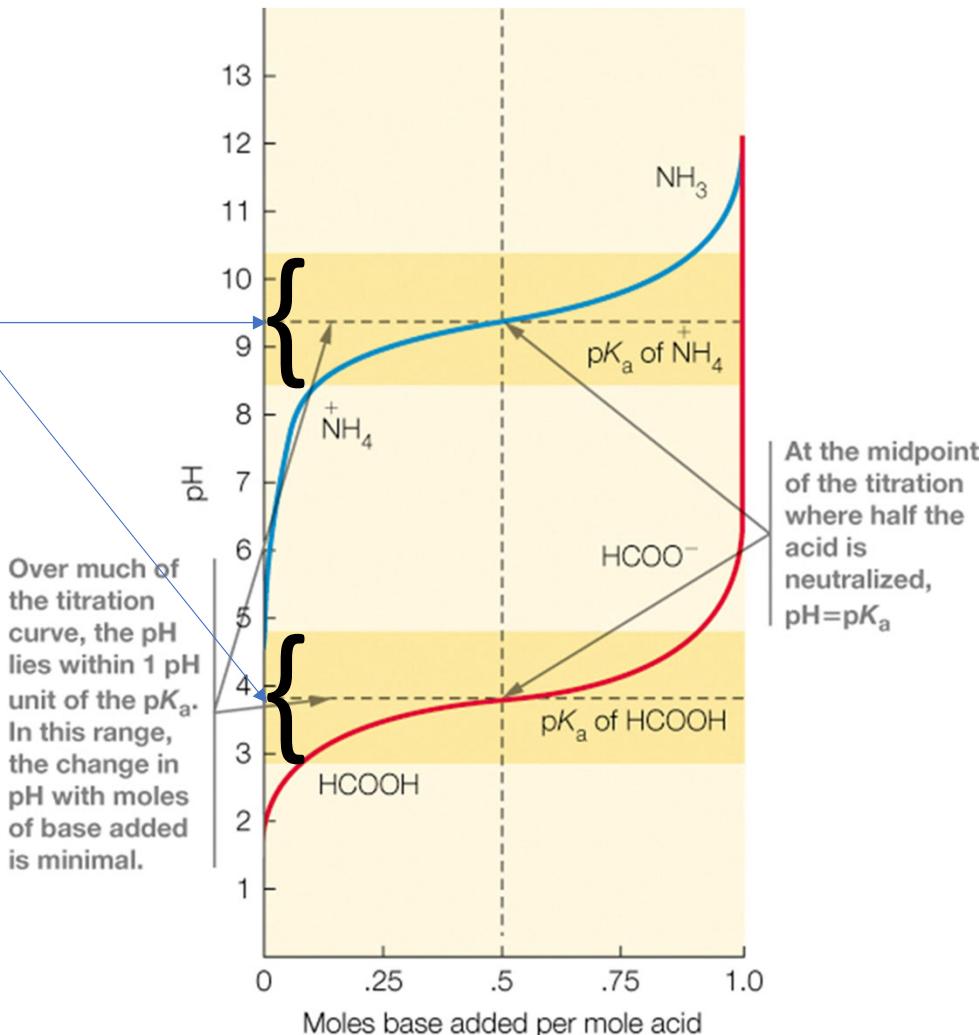


FIGURE 2.19 Titration curves of weak acids.

Buffer Solutions

- Buffers will **resist pH change following the addition of acid or base** within about ± 1 pH unit of the pK_a of a solution
- at pH values ± 3 pH units from pK_a , the group is essentially fully deprotonated or fully protonated
- **at pH = pK_a , the group is 50% protonated, thus it is 50% ionized**
- The Henderson–Hasselbalch equation can be used to calculate the average charge on an ionizable group at any pH

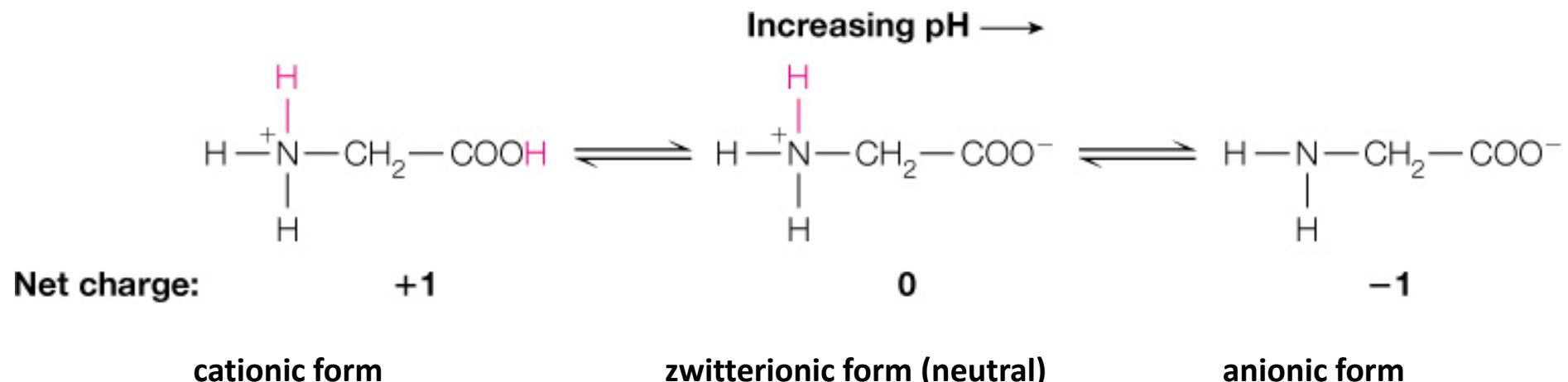


The pH of a buffer – An example calculation

- Consider an acetic acid buffer ($pK_a = 4.76$) containing 1 M HOAc and 1 M Na^+OAc^-
- $\text{pH} = pK_a + \log[1 \text{ M}/1 \text{ M}] = 4.76$
- Now add 0.1 mol of HCl to 1.0 liter of this buffer; H^+ reacts quantitatively with OAc^- to drive the pH down:
 - $[\text{HOAc}] = 1 \text{ M} + 0.1 \text{ M} = 1.1 \text{ M}$
 - $[\text{OAc}^-] = 1 \text{ M} - 0.1 \text{ M} = 0.9 \text{ M}$
- $\text{pH} = 4.76 + \log(0.9/1.1) = 4.76 - 0.087 = 4.673$

Amino Acids: Multiple Ionizing Groups

- An **ampholyte** is a molecule that contains groups with both acidic and basic pK_a values
 - many biologically relevant molecules are ampholytes
 - Amino acids are ampholytes: e.g. glycine (R group = H)



The neutral form is an identifiable characteristic

- The **isoelectric point (pI)** is the pH at which the average charge on the molecule is zero
- For a simple molecule, it is the average of the two pK_a values surrounding the isoelectric species
- For glycine:
 - $pI = (2.3 + 9.6)/2 = 5.95$
 - pK_a values can be determined by titration (Prac 1).

$$pI = \frac{pK_{COOH} + pK_{NH_3^+}}{2}$$

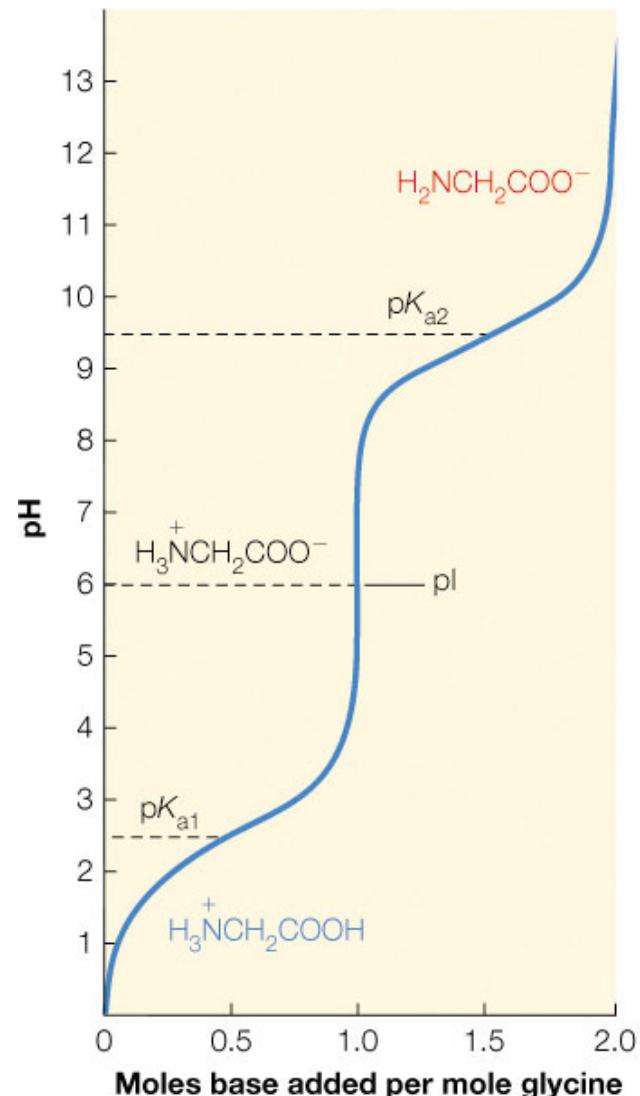


FIGURE 2.20 Titration of the amphotelyte glycine.

Which Group is Ionized?

Relative concentrations of glycine species are a function of pH

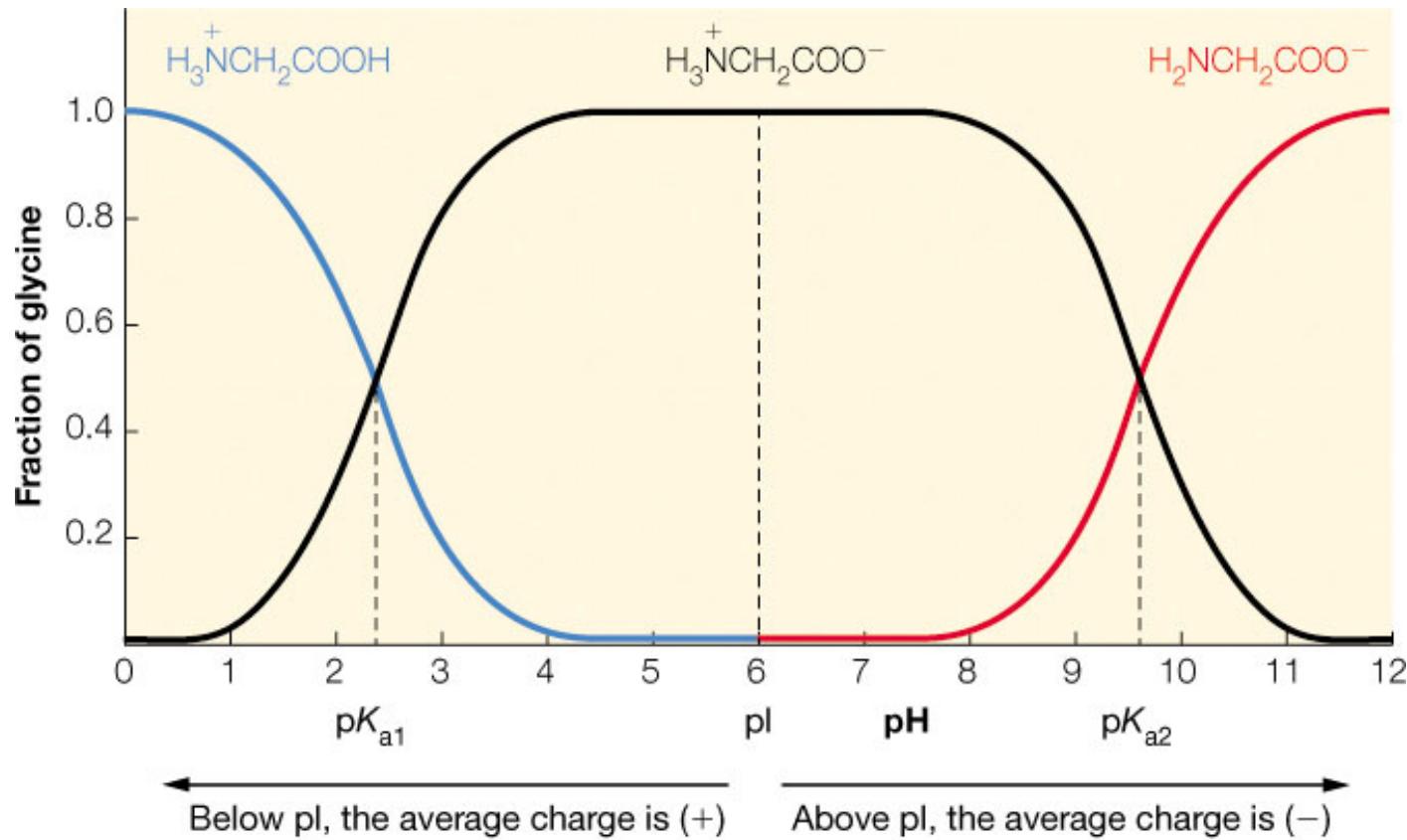


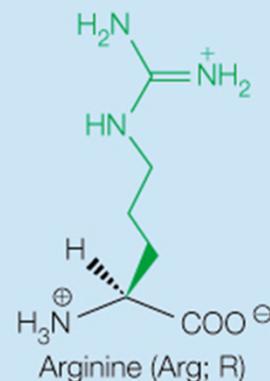
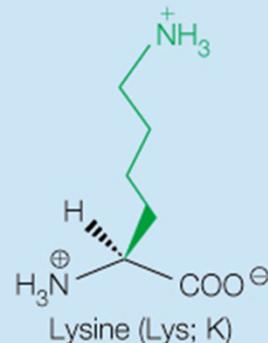
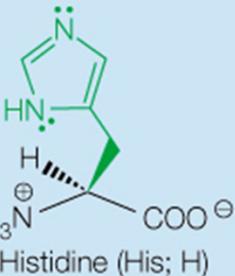
FIGURE 2.21 The relative concentrations of the three forms of glycine as a function of pH.

Five Amino Acids have R groups with Acid/Base properties

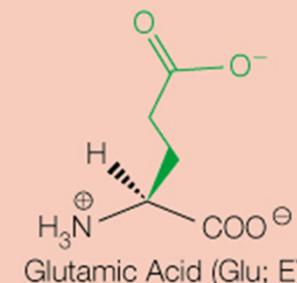
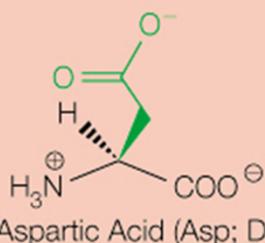
TABLE 5.1 Properties of the common amino acids found in proteins

| Name | Abbreviations: 1- and 3-letter codes | pK _a of α -COOH Group ^a | pK _a of α -NH ₃ ⁺ Group ^a | pK _a of Ionizing Side Chain ^a |
|---------------|--------------------------------------|--|--|---|
| Alanine | A, Ala | 2.3 | 9.7 | — |
| Arginine | R, Arg | 2.2 | 9.0 | 12.5 |
| Asparagine | N, Asn | 2.0 | 8.8 | — |
| Aspartic acid | D, Asp | 2.1 | 9.8 | 3.9 |
| Cysteine | C, Cys | 1.8 | 10.8 | 8.3 |
| Glutamine | Q, Gln | 2.2 | 9.1 | — |
| Glutamic acid | E, Glu | 2.2 | 9.7 | 4.2 |
| Glycine | G, Gly | 2.3 | 9.6 | — |
| Histidine | H, His | 1.8 | 9.2 | 6.0 |
| Isoleucine | I, Ile | 2.4 | 9.7 | — |
| Leucine | L, Leu | 2.4 | 9.6 | — |
| Lysine | K, Lys | 2.2 | 9.0 | 10.0 |

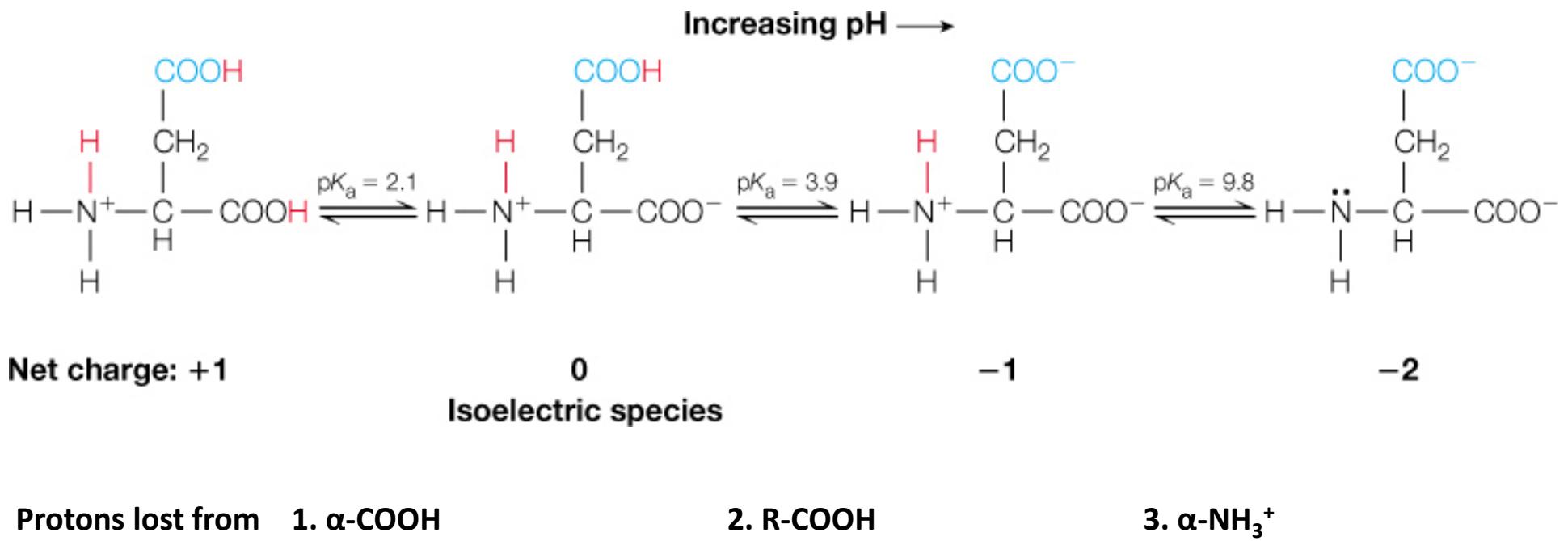
Positively Charged Polar Amino Acids



Negatively Charged Polar Amino Acids



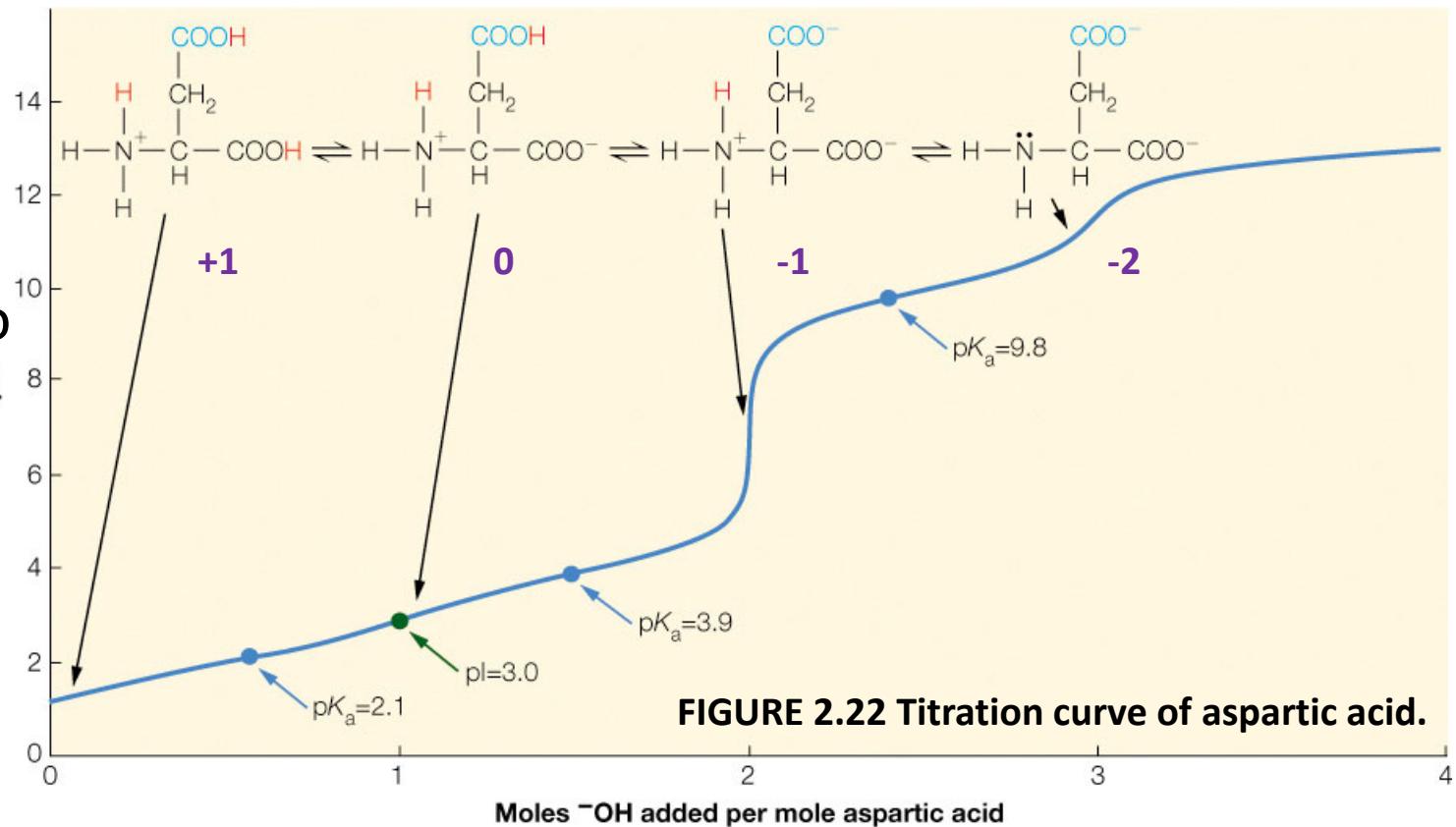
Aspartic Acid has three ionizing Groups



Aspartic Acid Titration Curve

pI = 3.0 (acidic!)

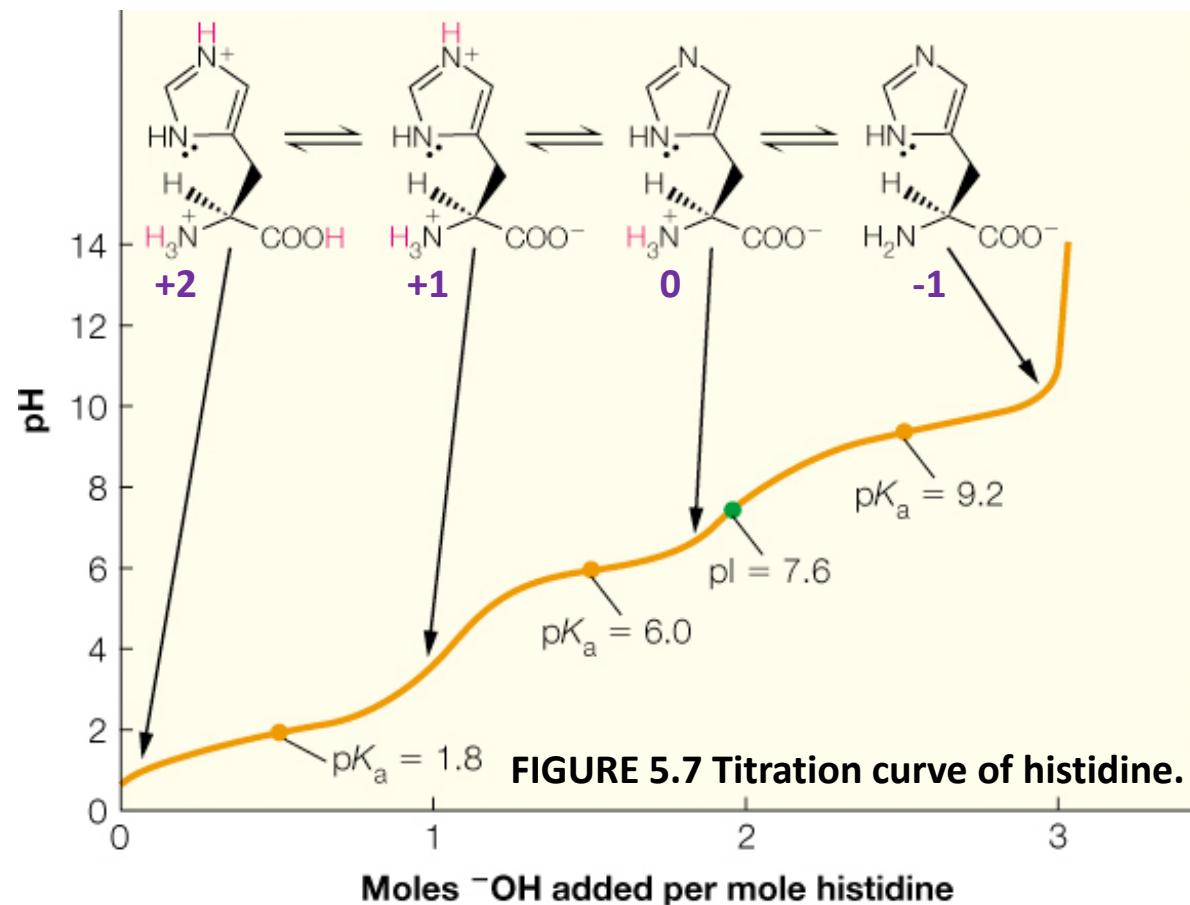
- The charge on aspartic acid (or aspartate) varies from +1 to -2, depending on pH
- Blue dots correspond to the pK_a values for the red hydrogen atoms
- The green dot corresponds to the pI (net charge on the molecule is zero)



Histidine is a base in neutral pH!

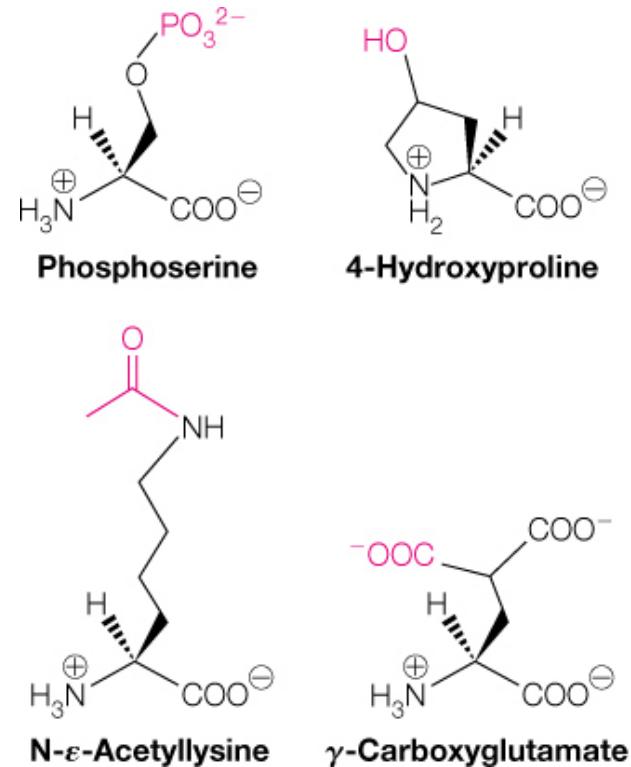
pI = 7.6

- The charge on histidine varies from +2 to -1, depending on pH
- Orange dots correspond to the pK_a values for the red hydrogen atoms
- The green dot corresponds to the pI (net charge on the molecule is zero)



Posttranslational Modification of Amino Acids

- Posttranslational modifications may function in signaling pathways, calcium binding, stabilizing structures such as collagen, or play roles in gene expression or suppression



Amino acids - summary

- The 20 standard amino acids share a common structure but differ in their side chains.
- Some amino acid side chains contain ionizable groups whose pK values may vary.
- The charge on amino acid side chains is retained in the proteins they make!

Acid-base properties of amino acids - Summary

- The 5 charged amino acid side chains contain ionizable groups with specific pK values.
 - Acidic side chains (Asp, Glu) have an extra carboxyl group
 - Basic side chains (Arg, Lys) have an extra basic group
- Histidine has a sidechain that can act as a base at physiological pH and is therefore considered a basic amino acid.
- The charges at the ends of the residues in a polypeptide or protein cancel out each other; but the charged side chains are charged at physiological pH.

(Theory for Practical 1)

L2 Spot test

- Now on iLearn
 - See Section 2 (Week 2)
- Available until this weekend (Sunday midnight)
- 2 questions; 5 mins; 1 attempt

Amino Acids 2: Physical Properties and Making Proteins from Genes

Shoba Ranganathan

Applied Biosciences

T: 02 9850 6262; E: shoba.ranganathan@mq.edu.au

Objectives

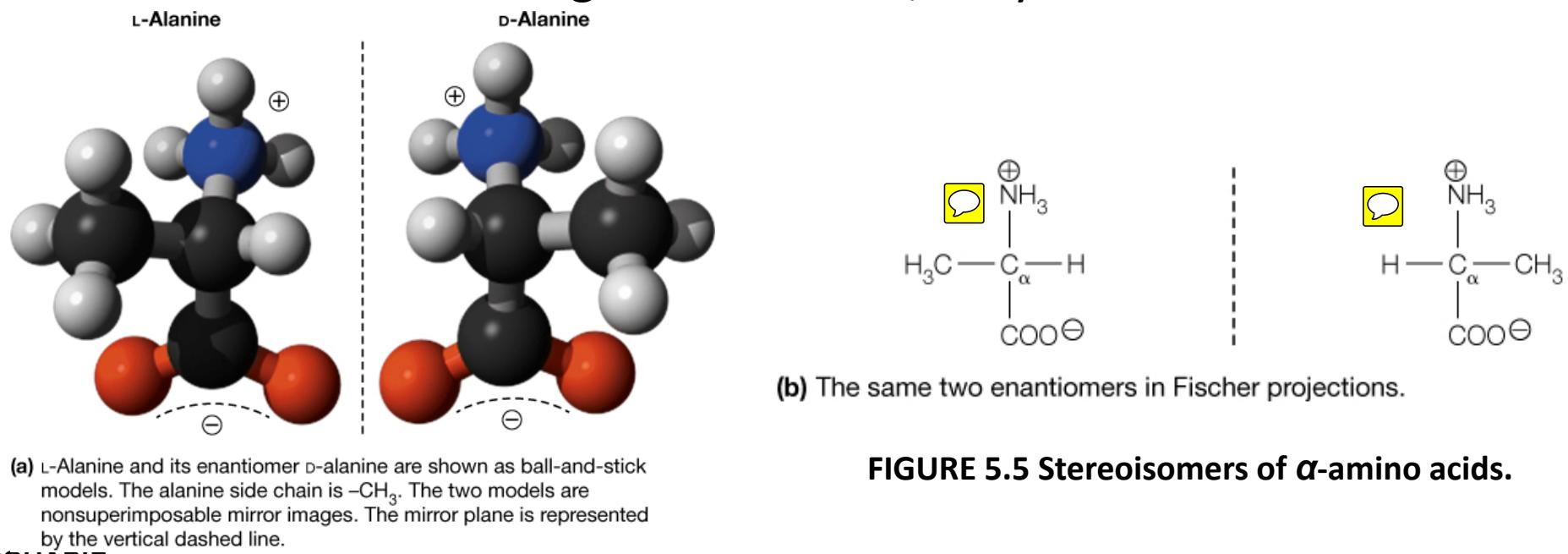
- Chirality ('handedness') in amino acids
- Absorption of light
- Covalent modification of amino acid side chains
- Biological function of amino acids or their derivatives
- How are proteins assembled from amino acids?

AAM: Chapters 5, 6, 4 and 1



α -Amino Acid Stereochemistry - 2

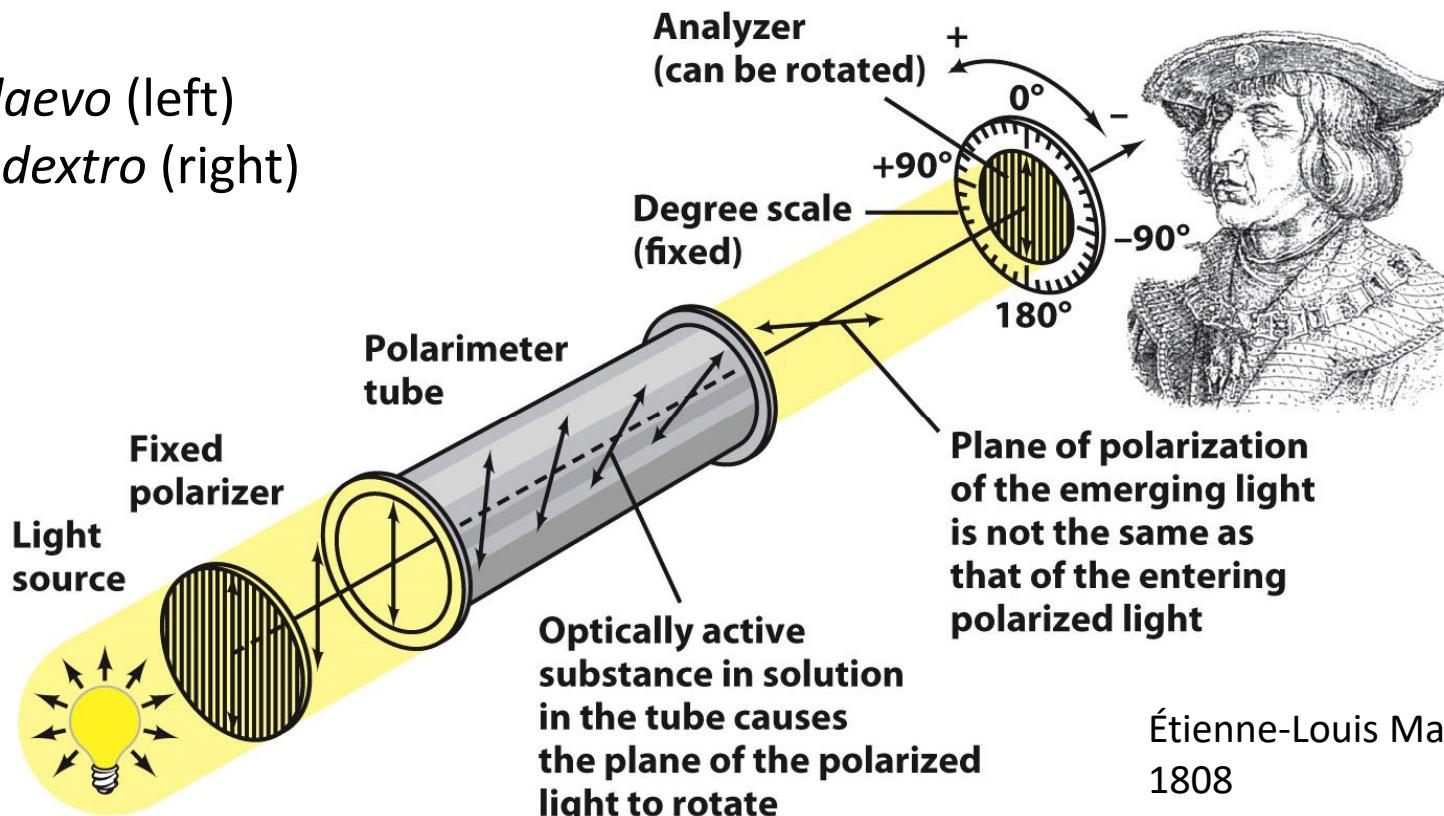
- All α -amino acids (20), except glycine, contain an asymmetric α -carbon
- When R = -CH₃ (methyl group), we have the amino acid alanine
- L-alanine is the mirror image of D-alanine; they are enantiomers



Chiral molecules rotate polarized light

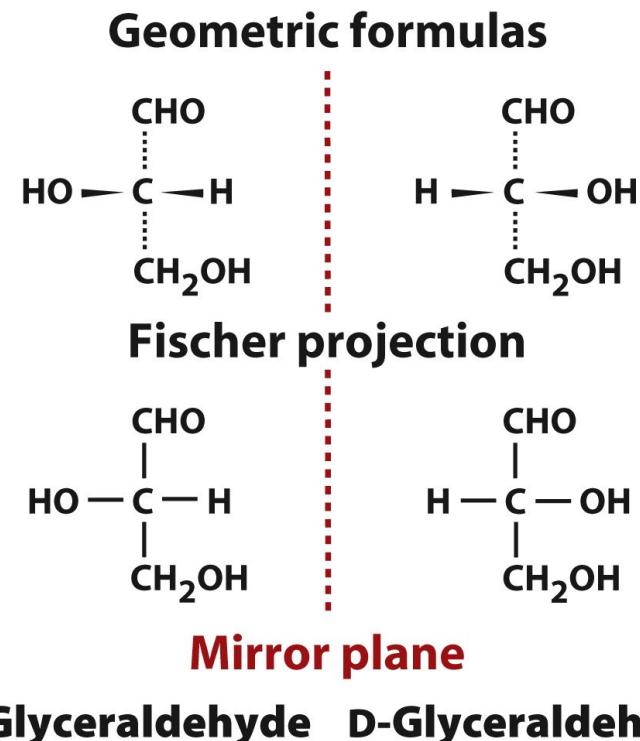
L: *laevo* (left)

D: *dextro* (right)

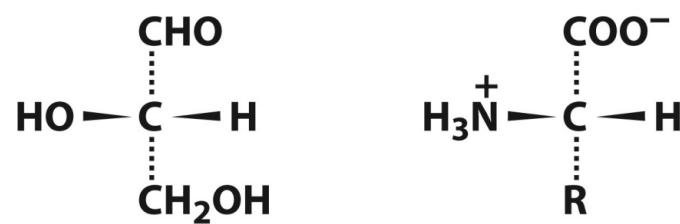


Étienne-Louis Malus,
1808

Humans have all L-amino acids



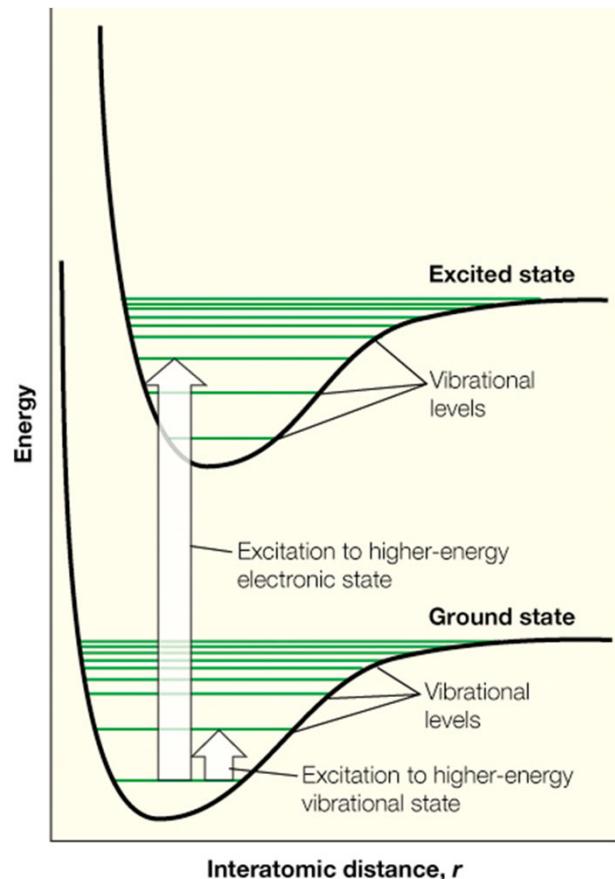
Relative configuration



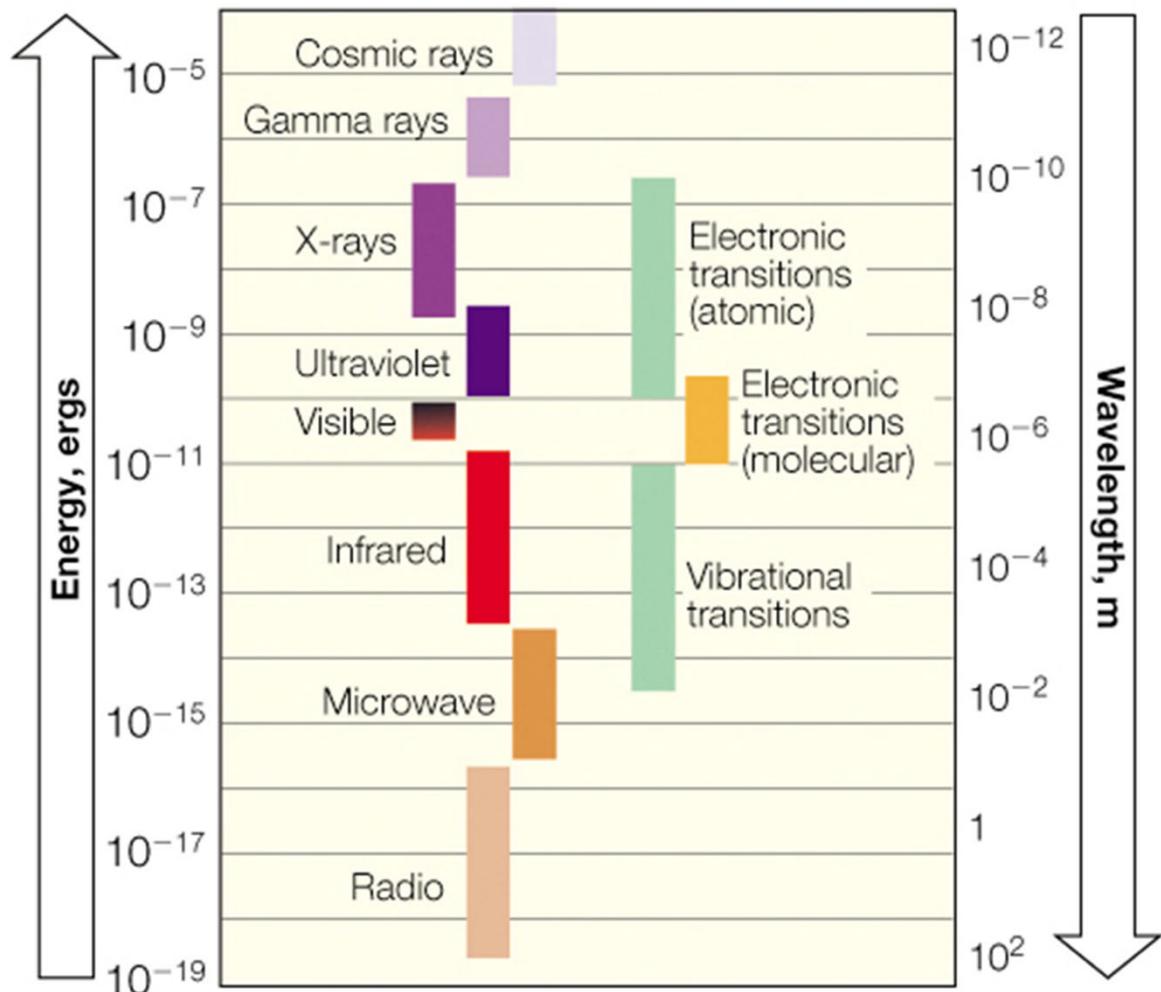
L-Glyceraldehyde L- α -Amino acid

- Organic synthesis invariably leads to a mixture of D- and L- isomers (racemic mixture)
 - Ibuprofen is marketed as a mixture, although the D-isomer (*S*-) has greater activity!
- However, biochemical reactions result in 100% of one stereo isomer!

Amino acids absorb light



(a) Electronic and vibrational transitions between allowed states in a diatomic molecule.



(b) The electromagnetic spectrum.

Figure 6A. The principles of absorption spectroscopy



UV absorption spectra of tyrosine and tryptophan

- Proteins can be quantified by examining the UV absorbance at 280 nm, which is dominated by tyrosine (Tyr) and tryptophan (Trp)
- In comparison, nucleic acids absorb most strongly at 260 nm

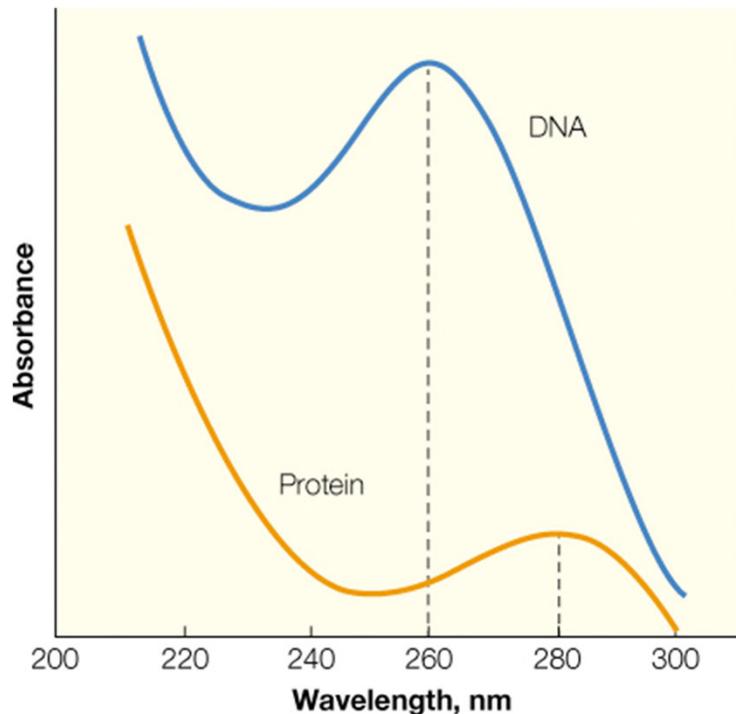
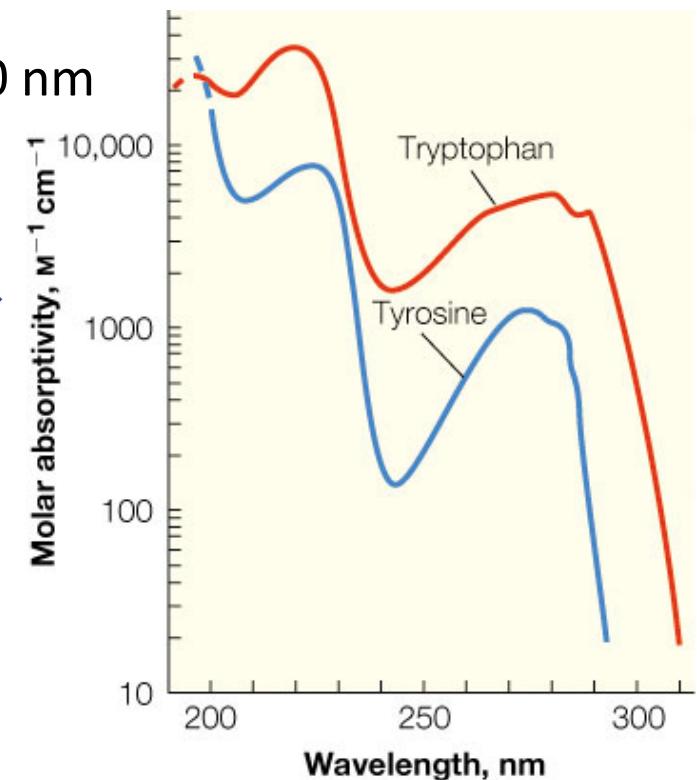


FIGURE 5.6 Absorption spectra of two aromatic amino acids in the near-ultraviolet region.

Figure 6A.2, The principles of absorption spectroscopy.



Spectrophotometer for measuring light absorption

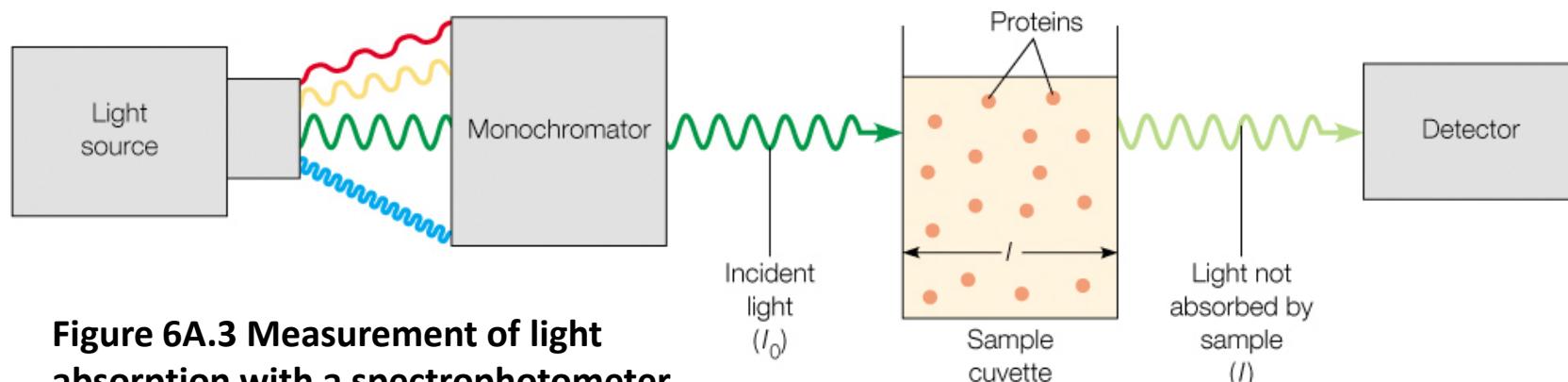


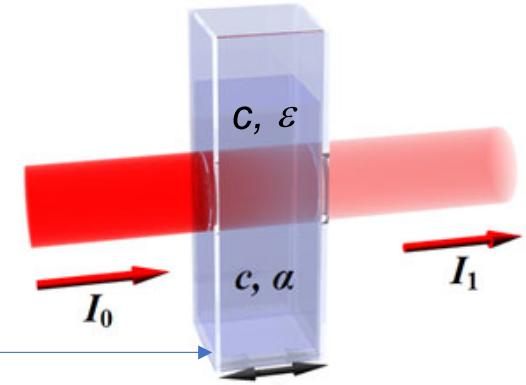
Figure 6A.3 Measurement of light absorption with a spectrophotometer.

- We will mainly be dealing with absorbance of light in the **visible region** (800 nm - 350 nm) and the **upper end of the ultraviolet region** (350 nm – 250 nm) of the spectrum
 - Many biological molecules absorb light in this range
 - It is thus a useful tool in examining biochemical processes
 - Instrumentation is relatively easy to use and the sample is placed in a cuvette
- Applications
 - identification by characteristic frequency of absorbance
 - amount of protein concentration present e.g. creatine (a protein) in urine.

Beer-Lambert's Law (aka Lambert-Beer's or Beer's Law)

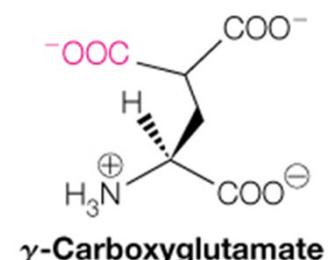
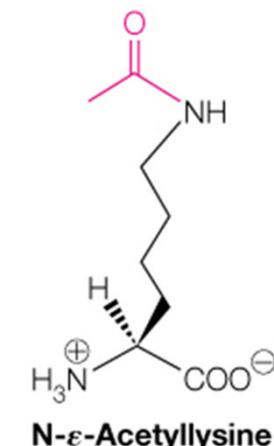
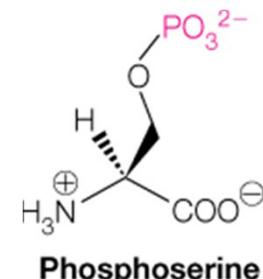
$$\log \frac{I_0}{I} = \epsilon \cdot c = A \text{ (what we measure, in a cuvette)}$$

- where
 - I/I_0 is the fraction of light absorbed - this is also called transmittance and the absorbance $A = -\log T = -\log (I/I_0)$
 - ϵ (Greek epsilon) is the molar extinction coefficient, which is a measure of absorption in $M^{-1} cm^{-1}$ or $L mol^{-1} cm^{-1}$
 - c is the pathlength of light passing through the sample in cm (usually 1 cm)
 - c is the concentration in molar units: **M** or **mol L⁻¹**
 - A is called the absorbance (ratio – so no units)
 - Valid for dilute solutions
- ❖ Light absorption is used to **quantitatively** estimate the **amount of protein** in a biological sample (**Prac 2**).
- ❖ This property can also be used to **identify** amino acids (and proteins) and nucleotide bases in a biological sample (**Prac 4**).



Protein side chains can be modified

- After synthesis (i.e. post-translational), some protein side chains can be modified usually by simple chemical reactions:
 - Hydroxylation, methylation, acetylation, phosphorylation, carboxylation, glycosylation
 - N- and C- ends can also be modified or blocked.
- Some modifications are essential for biological function.
 - signaling pathways, calcium binding, stabilizing structures such as collagen, or play roles in gene expression or suppression
- Many side chains can be modified in a protein.



N- and C-termini blocking and Disulfide bridges

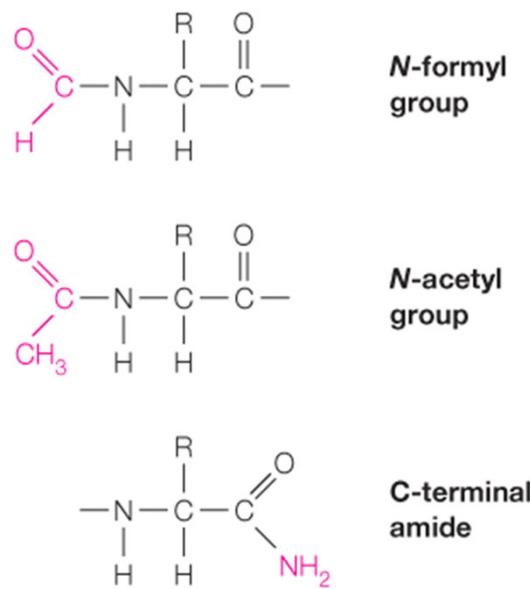
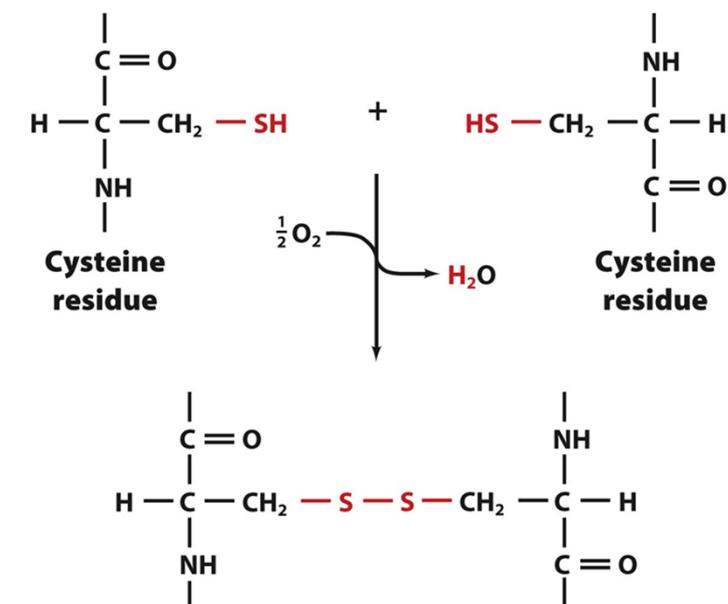


FIGURE 5.12 Groups that may block N- or C-termini in proteins.

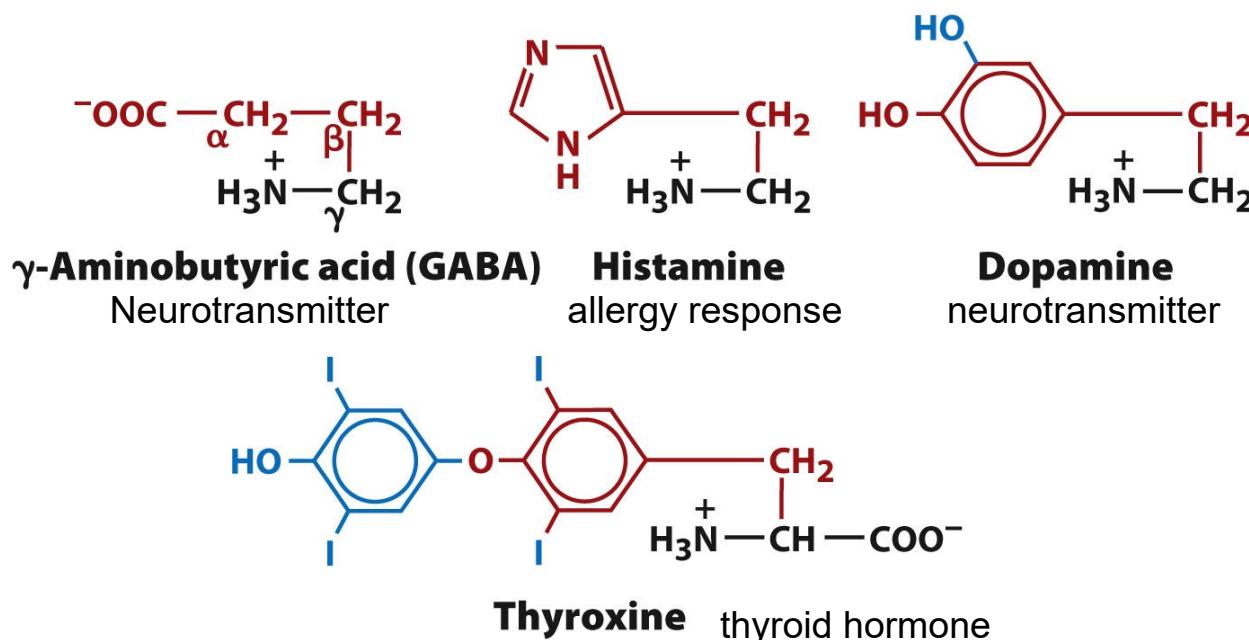
Side chains of **cysteine residues** can get oxidized to form a disulphide($-\text{S}-\text{S}-$) bond.

- Link different parts of the same protein or even different protein chains.
- Increase protein stability.



Some amino acids are biologically active

- Some amino acids transport nitrogen while others can provide energy as a fuel source (amino acid metabolism: later lectures)
- Some function as chemical messengers for cellular communication: e.g. **glycine** and **glutamate** are neurotransmitters
- Example **biologically active amino acid derivates** are shown below:



Peptide Bond Formation between Amino Acids

- The **condensation** of two amino acids forms a **peptide bond** and releases **water**
 - The charges in the middle are lost!
 - The ends are still charged.
- This reaction does not occur spontaneously, and is coupled to ATP hydrolysis during protein biosynthesis
 - Can be extended from the carboxyl end.....

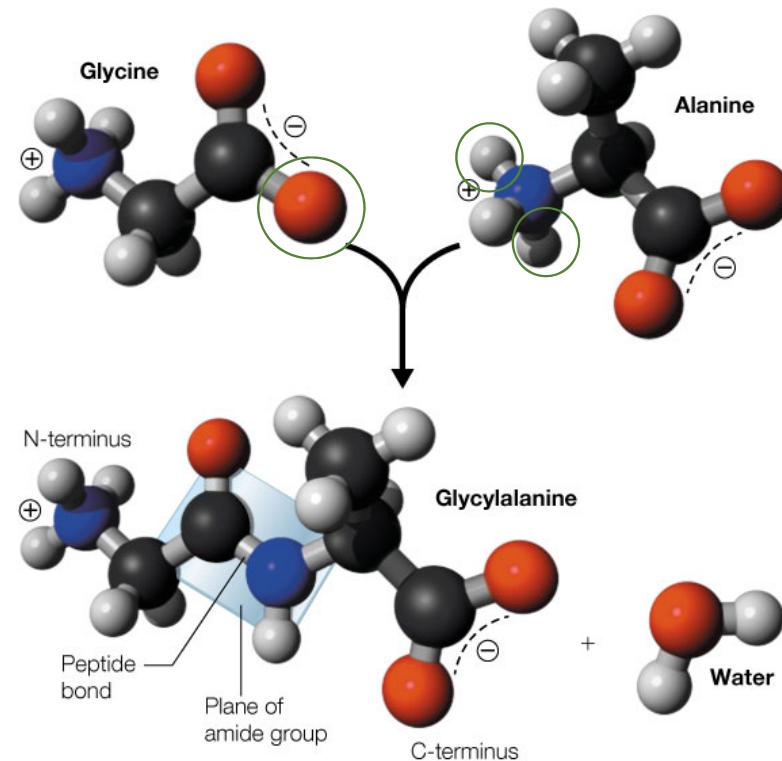


FIGURE 5.9 A peptide bond between amino acids.

Example peptides

- Important parts of a peptide, IFDTH
- The main chain is also called the backbone
- The ends are normally charged
- Amino acids in a peptide or protein are called **residues**

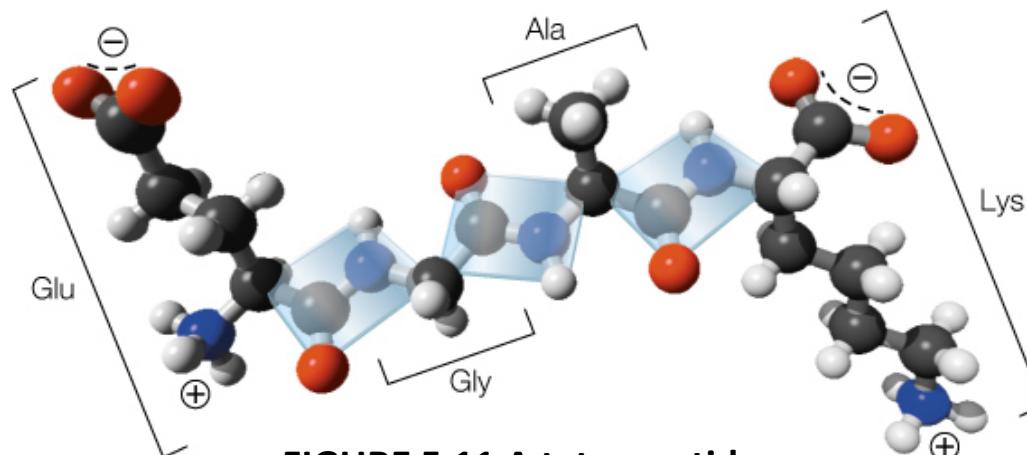
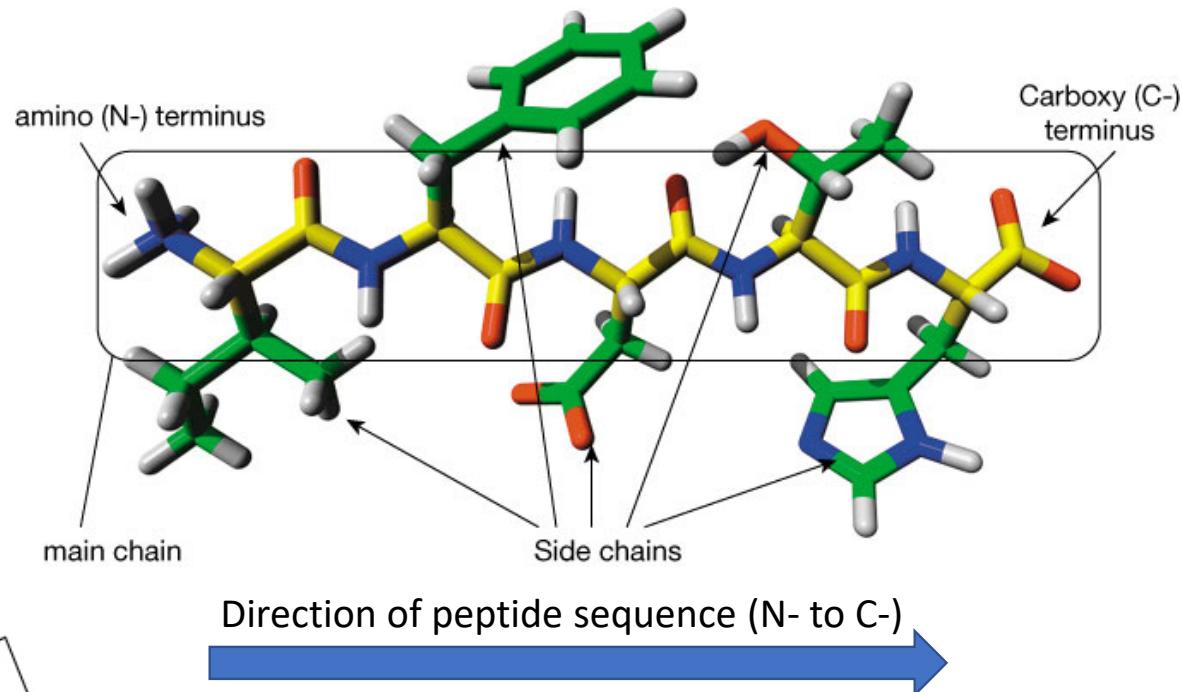
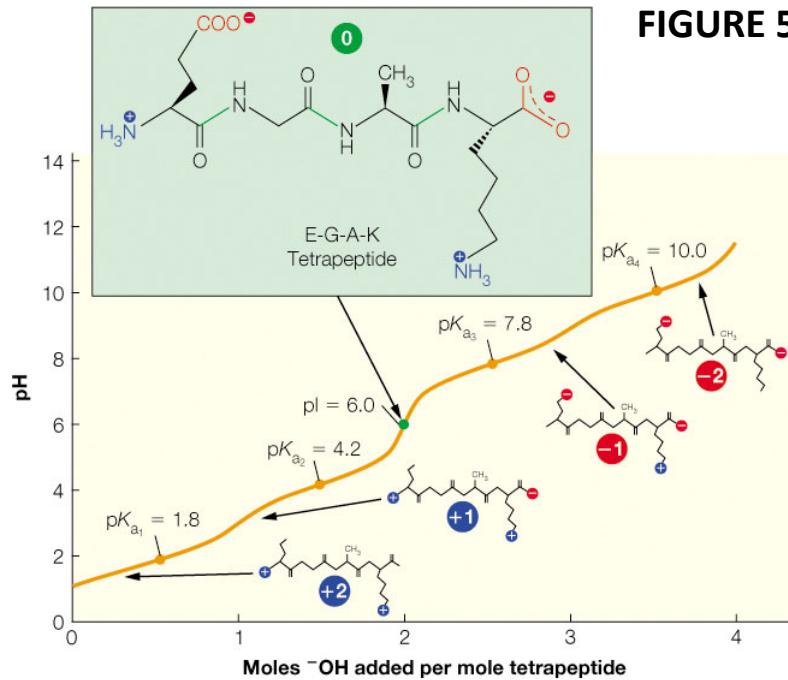


FIGURE 5.11 A tetrapeptide.



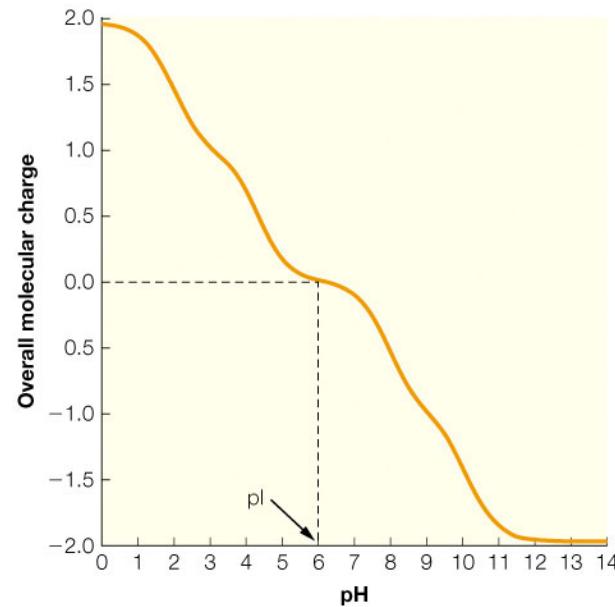
- Tetrapeptide:
- Glu-Gly-Ala-Lys
- EGAK

Peptides and Proteins as Polyampholytes



(a) This titration curve for the tetrapeptide Glu-Gly-Ala-Lys shows the major ionization states present as a function of pH. The tetrapeptide is shown schematically with (+) charges in blue and (−) charges in red. Net charges for the different ionization states are shown in solid circles.

FIGURE 5.14 Ionization behavior of a tetrapeptide.



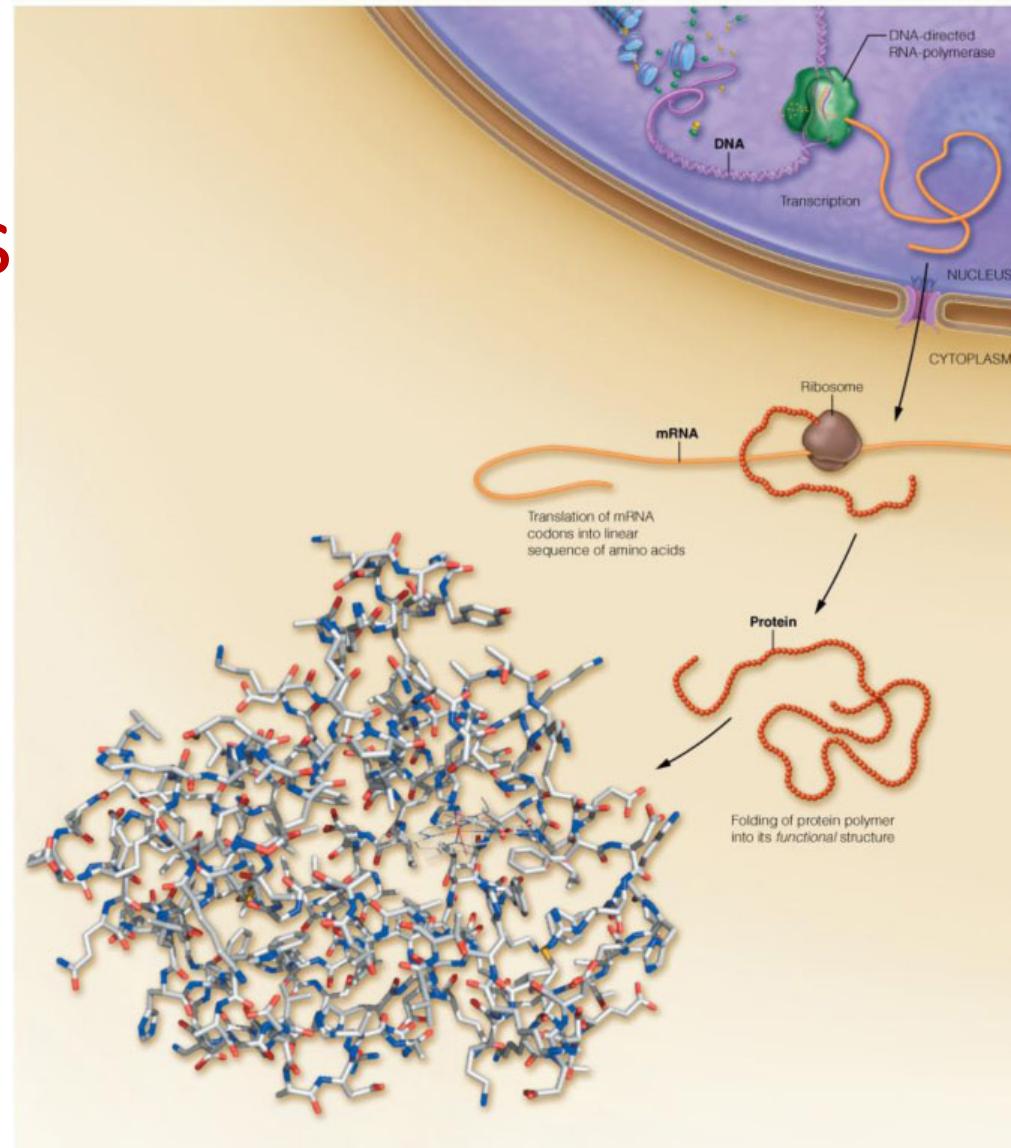
(b) The net charge on Glu-Gly-Ala-Lys as a function of pH is not a step function. Rather, between pH 1 and 11, it changes gradually and smoothly from +2 to −2. At values of pH > pl, the overall charge on the tetrapeptide is negative. At values of pH < pl, the net charge is positive.

As pH increases the overall charge on a peptide will become more negative; as pH decreases it will become more positive.



Amino acids are the building blocks of proteins

- DNA is transcribed to form messenger RNA (mRNA) in the nucleus
- mRNA is exported to the cytoplasm, where it is bound to the ribosome
- mRNA is translated into **a linear sequence of amino acids** that folds into a 3D structure



From amino acid monomers to polymers

Peptides

- Short polymers of amino acids (aa)
- Each unit is called a residue
- 2 residues - dipeptide
- 3 residues - tripeptide
- 12-20 residues - oligopeptide
- many - polypeptide

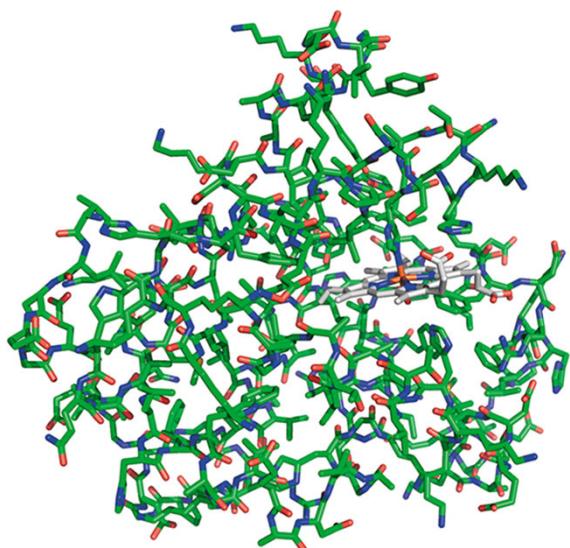
Proteins

- 1 or more chains of at least 40 aa
- One polypeptide chain - a monomeric protein
- More than one - multimeric protein
- Homomultimer - one kind of chain
- Heteromultimer - two or more different chains
- Hemoglobin, for example, is a heterotetramer, with α (alpha) chains and β (beta) chains



Proteins: polypeptides with Defined Sequences

- Every protein has a defined number and order of amino acids.
- This is called its “primary structure”
- The sequence of the sperm whale myoglobin determines its final biological form (folded) and its function

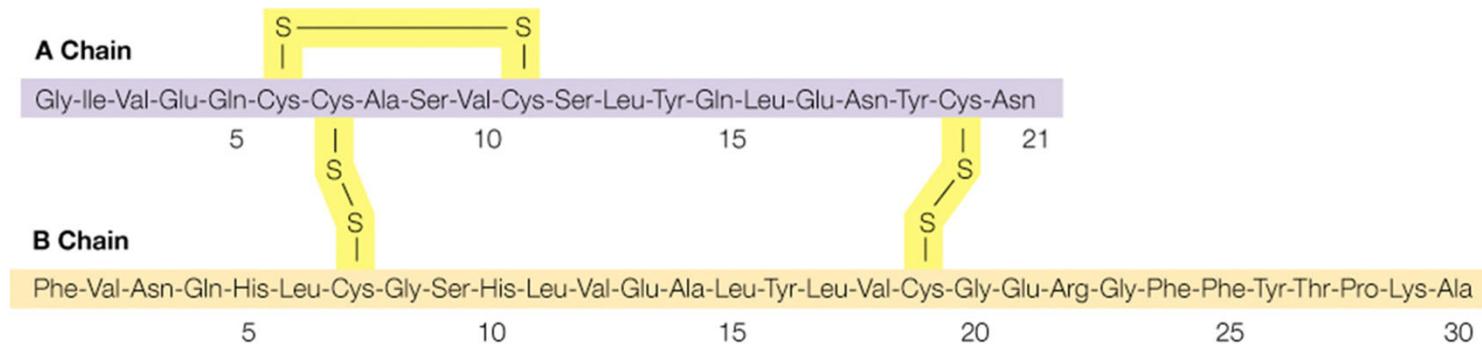


- Its similarity to the human myoglobin sequence is remarkable!
- Of 153 amino acids, 128 (84% are identical)
- 13 residues (green) have chemically similar side chains – adding these, 92% are similar!

| Number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|--------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Human | G | L | S | D | G | E | W | Q | L | V | L | N | V | W | G |
| Whale | V | L | S | E | G | E | W | Q | L | V | L | H | V | W | A |
| Number | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
| Human | K | V | E | A | D | I | P | G | H | G | Q | E | V | L | I |
| Whale | K | V | E | A | D | V | A | G | H | G | Q | D | I | L | I |
| Number | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 | 41 | 42 | 43 | 44 | 45 |
| Human | R | L | F | K | G | H | P | E | T | L | E | K | F | D | K |
| Whale | R | L | F | K | S | H | P | E | T | L | E | K | F | D | R |
| Number | 46 | 47 | 48 | 49 | 50 | 51 | 52 | 53 | 54 | 55 | 56 | 57 | 58 | 59 | 60 |
| Human | F | K | H | L | K | S | E | D | E | M | K | A | S | E | D |
| Whale | F | K | H | L | K | T | E | A | E | M | K | A | S | E | D |
| Number | 61 | 62 | 63 | 64 | 65 | 66 | 67 | 68 | 69 | 70 | 71 | 72 | 73 | 74 | 75 |
| Human | L | K | K | H | G | A | T | V | L | T | A | L | G | G | I |
| Whale | L | K | K | H | G | V | T | V | L | T | A | L | G | A | I |
| Number | 76 | 77 | 78 | 79 | 80 | 81 | 82 | 83 | 84 | 85 | 86 | 87 | 88 | 89 | 90 |
| Human | L | K | K | K | G | H | H | E | A | E | I | K | P | L | A |
| Whale | L | K | K | K | G | H | H | E | A | E | L | K | P | L | A |
| Number | 91 | 92 | 93 | 94 | 95 | 96 | 97 | 98 | 99 | 100 | 101 | 102 | 103 | 104 | 105 |
| Human | Q | S | H | A | T | K | H | K | I | P | V | K | Y | L | E |
| Whale | Q | S | H | A | T | K | H | K | I | P | I | K | Y | L | E |
| Number | 106 | 107 | 108 | 109 | 110 | 111 | 112 | 113 | 114 | 115 | 116 | 117 | 118 | 119 | 120 |
| Human | F | I | S | E | C | I | I | Q | V | L | Q | S | K | H | P |
| Whale | F | I | S | E | A | I | I | H | V | L | H | S | R | H | P |
| Number | 121 | 122 | 123 | 124 | 125 | 126 | 127 | 128 | 129 | 130 | 131 | 132 | 133 | 134 | 135 |
| Human | G | D | F | G | A | D | A | Q | G | A | M | N | K | A | L |
| Whale | G | N | F | G | A | D | A | Q | G | A | M | N | K | A | L |
| Number | 136 | 137 | 138 | 139 | 140 | 141 | 142 | 143 | 144 | 145 | 146 | 147 | 148 | 149 | 150 |
| Human | E | L | F | R | K | D | M | A | S | N | Y | K | E | L | G |
| Whale | E | L | F | R | K | D | I | A | A | K | Y | K | E | L | G |
| | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | |

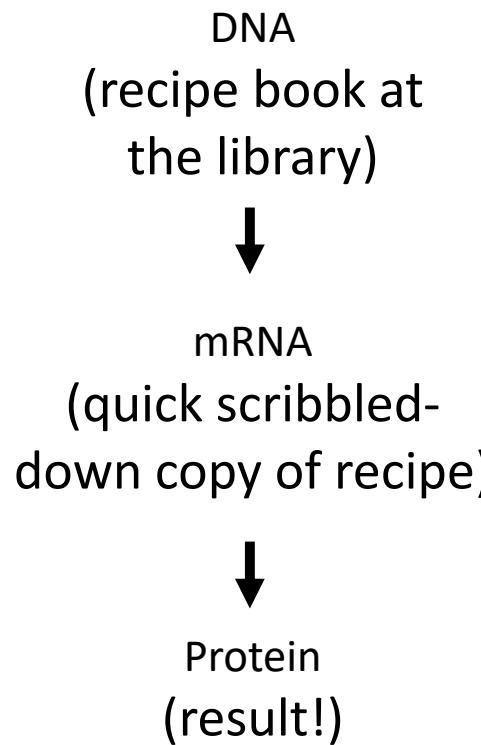
How small or large are proteins?

- Insulin (bovine) – a small protein with an A chain of 21 residues and a B chain of 30 residues - total mol. wt. of 5,733



- Largest known protein is **titin** in muscle fibres: 34,350 residues (MW 3816 kD)
- Majority of proteins contain 100-1000 aa.
- Multisubunit proteins: many identical or non-identical polypeptide chains called subunits
 - Insulin has two chains.
 - There are two disulfide bonds holding the chains together

From Genes to Proteins



| | U | C | A | G | |
|---|--|------------------------------|--|---------------------------------------|------------------|
| U | UUU Phe UUC UUA UUG | UCU Ser UCC UCA UCG | UAU Tyr UAC UAA Stop UAG Stop | UGU Cys UGC UGA Stop UGG Trp | U C A G |
| C | CUU Leu CUC CUA CUG | CCU Pro CCC CCA CCG | CAU His CAC CAA Gln CAG | CGU Arg CGC CGA CGG | U C A G |
| A | AUU Ile AUC AUA AUG Met/ start | ACU Thr ACC ACA ACG | AAU Asn AAC AAA Lys AAG | AGU Ser AGC AGA Arg AGG | U C A G |
| G | GUU Val GUC GUA GUG | GCU Ala GCC GCA GCG | GAU Asp GAC GAA Glu GAG | GGU Gly GGC GGA GGG | U C A G |

- The standard genetic code shows the **codons**, or **base triplets**, that correspond to each amino acid residue
 - A single amino acid can be encoded by 1-6 codons!
- The **AUG** start codon places a **Met** at the beginning of **eukaryotic proteins**,
- Three codons, **UAA**, **UGA**, and **UAG**, are stop codons, which signal termination of translation

Nucleic Acids make polymeric DNA

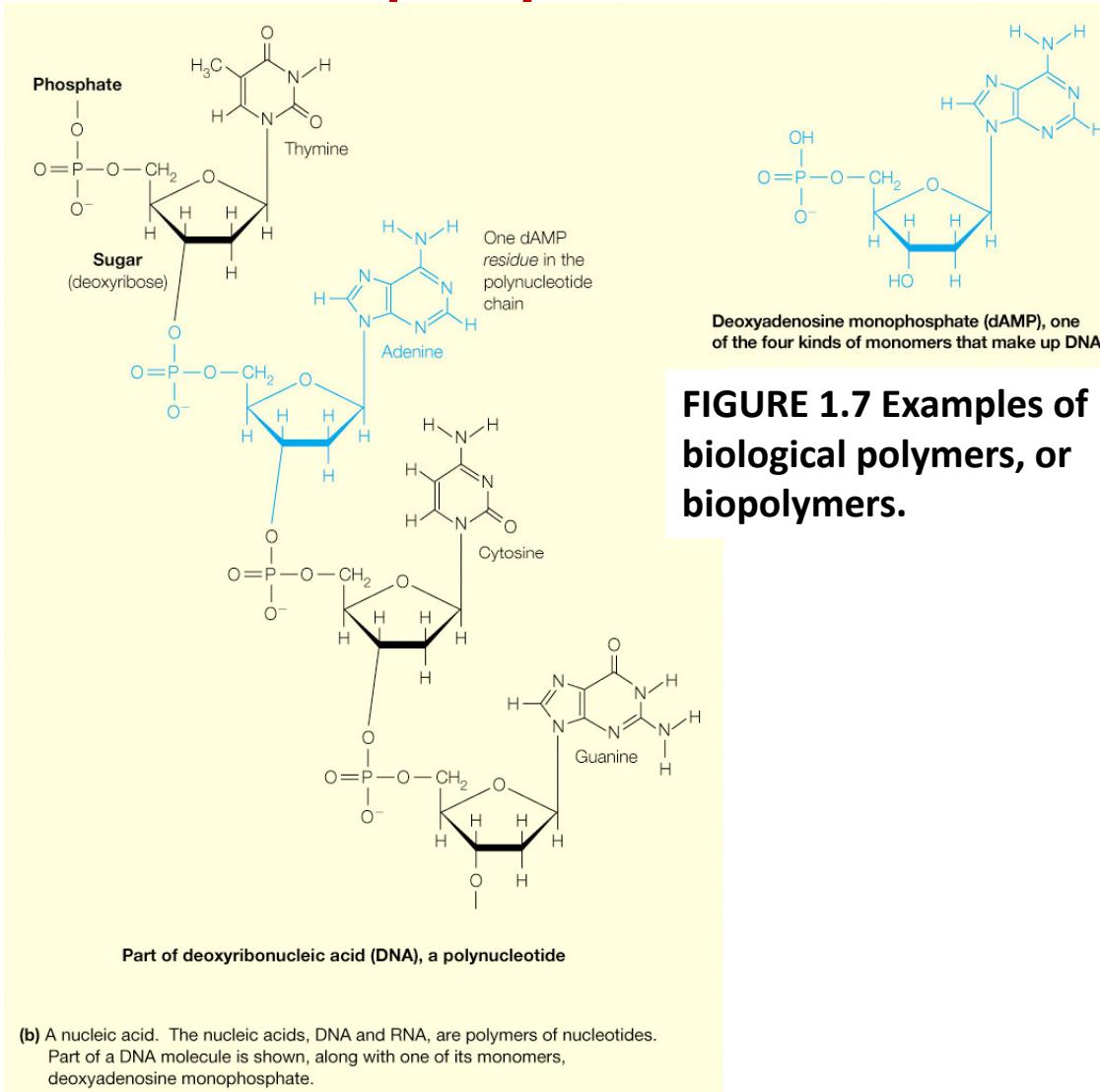


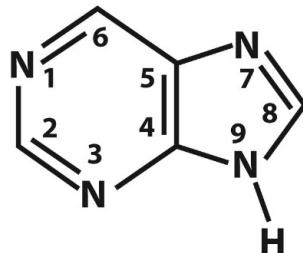
FIGURE 1.7 Examples of biological polymers, or biopolymers.

Nucleotide bases in nucleic acids

2 purines and 3 pyrimidines

Purines

- A: Adenine
- G: Guanine



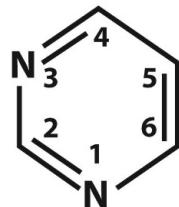
Purine

Unnumbered 3 p41a
© 2013 John Wiley & Sons, Inc. All rights reserved.

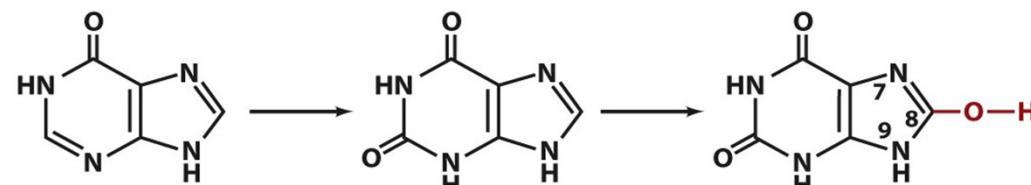
Other naturally occurring bases:

Pyrimidines

- C: Cytosine
- U: Uracil (RNA)
- T: Thymine (DNA)

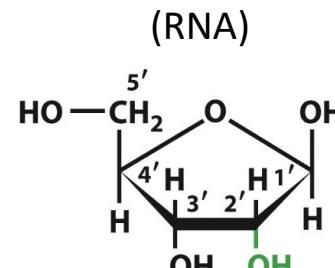


Pyrimidine

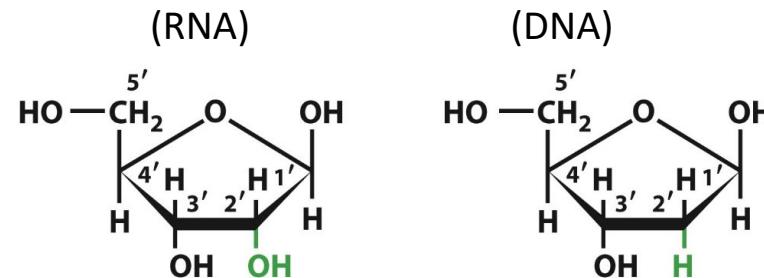


Hypoxanthine

Pentose sugars



Ribose

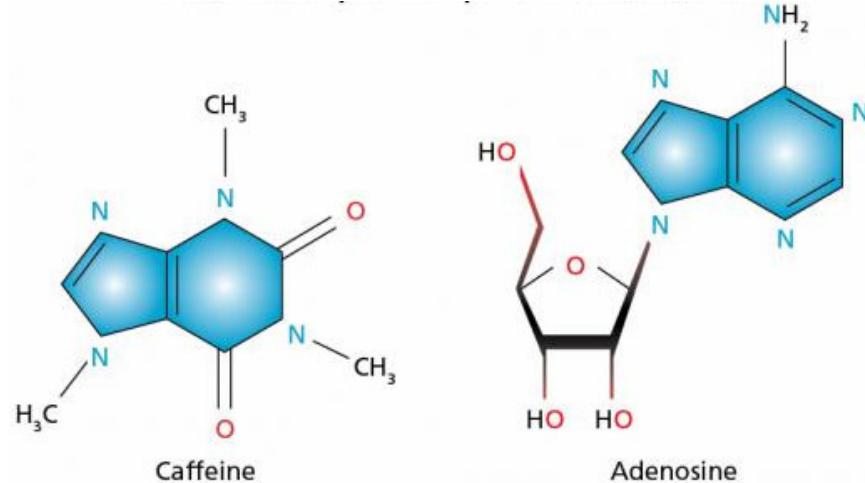


Deoxyribose

Unnumbered 3 p41b
© 2013 John Wiley & Sons, Inc. All rights reserved.

Module 1: Building Blocks of Biochemistry

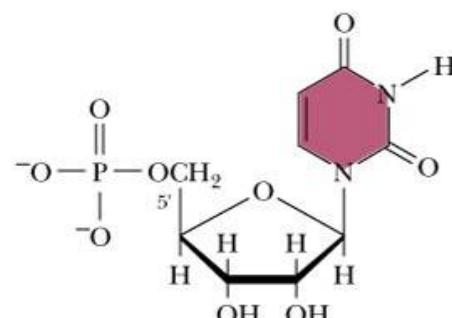
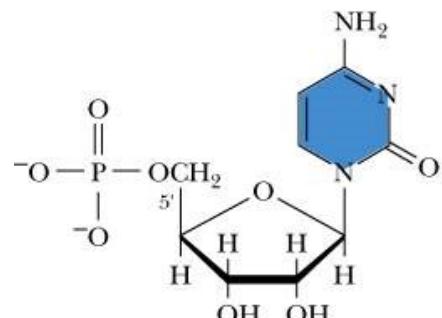
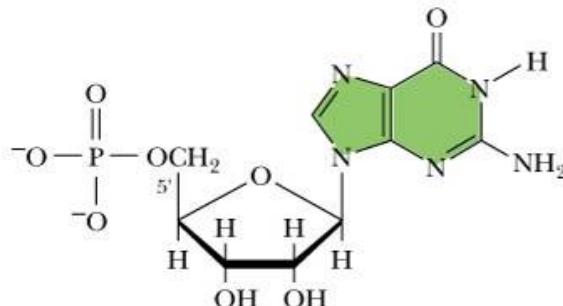
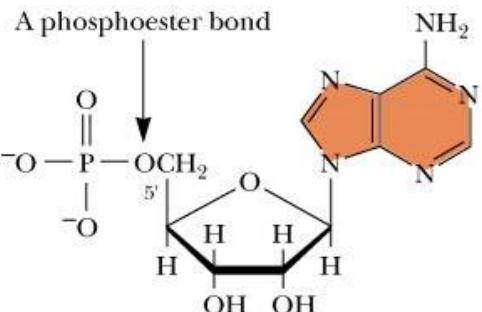
Adenosine: (adenine + sugar) has physiological activity



- is an **autacoid**, or local hormone, and neuromodulator.
- influences blood vessel dilation, smooth muscle contraction, neurotransmitter release, and fat metabolism.
- also a sleep regulator:
 - Adenosine rises during wakefulness, promoting eventual sleepiness.
- **Caffeine** (structurally similar to adenosine) promotes wakefulness by blocking binding of adenosine to its neuronal receptors (inhibitory action).

Base + ribose sugar + phosphate = nucleotide

Adenylate, guanylate, cytidylate, uridylate formed

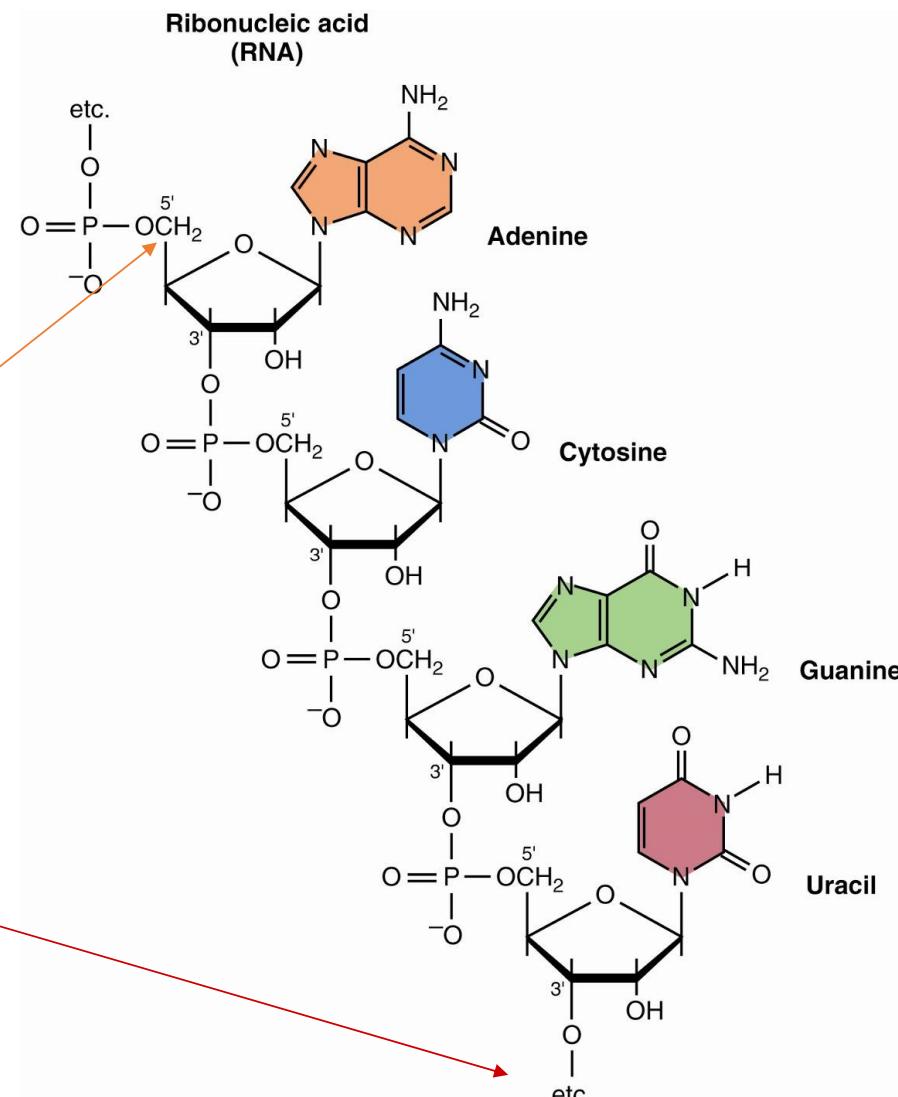


- Energy carriers with 2 or 3 phosphate groups, e.g. ATP
- All tri-phosphates are important for metabolism
- Cyclized forms can be messengers



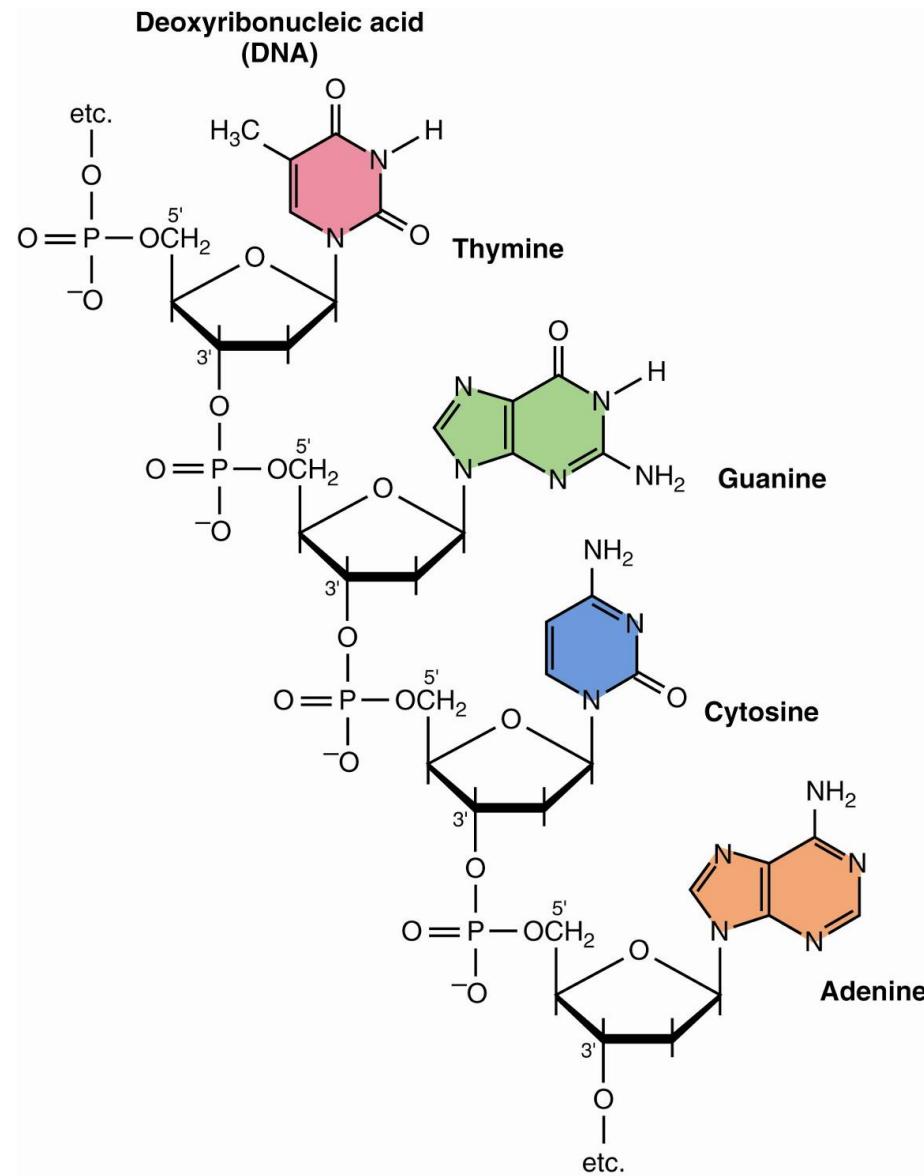
Making RNA from nucleotides

- Phosphodiester bridges link nucleotides together to form polynucleotide chains.
- The 5'-ends of the chains are at the top; the 3'-ends are at the bottom.
- Sequence is: ACGU



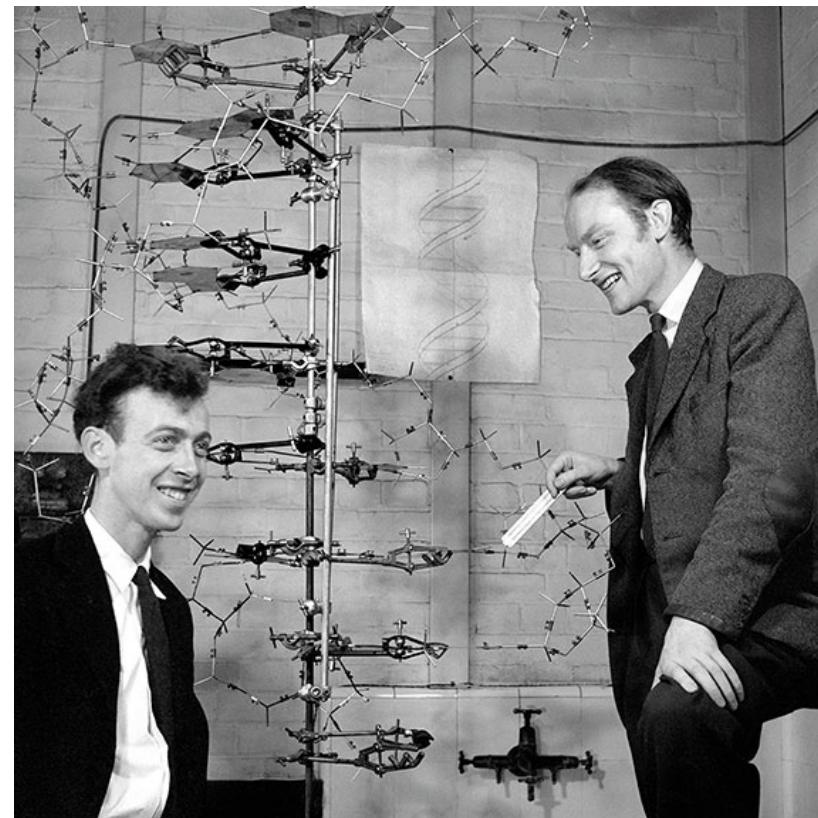
Making DNA from nucleotides

- Poly(deoxy)- nucleotide chains.
- The 5'-ends of the chains are at the top; the 3'-ends are at the bottom.
- Sequence is: TGCA
- DNA cannot be linked at 2' (deoxy)
- A pairs with T and G with C to form an anti-parallel double helix.
- A% = T% and G% = C%

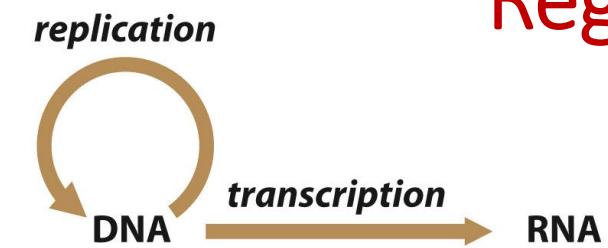


DNA structure solved in 1953

- The American biologist James Watson and the English physicist Francis Crick described the double-helical structure of DNA



Regions of DNA called Genes direct protein synthesis



Unnumbered 3 p50

© 2013 John Wiley & Sons, Inc. All rights reserved.

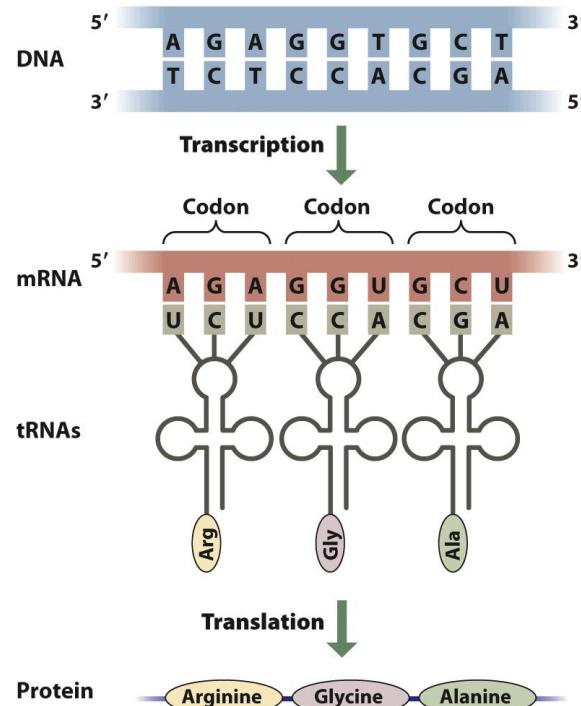


Figure 3-12
© 2013 John Wiley & Sons, Inc. All rights reserved.

Translation occurs in a special organelle called the ribosome

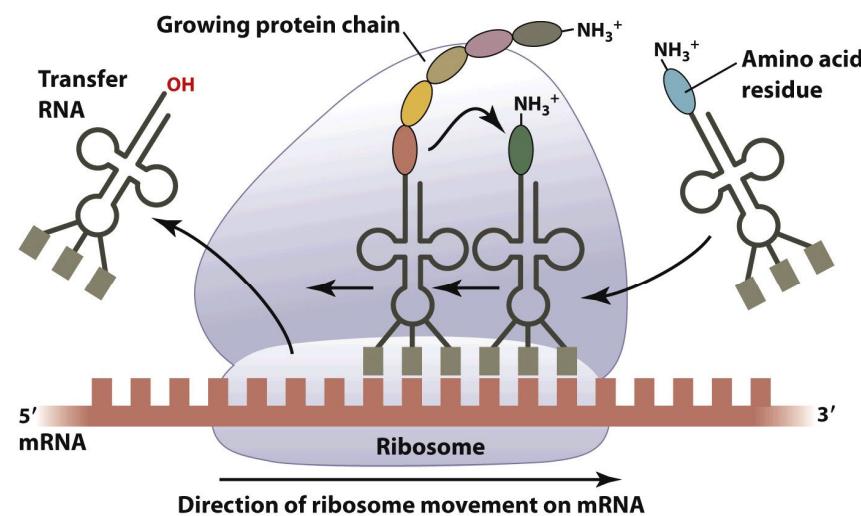
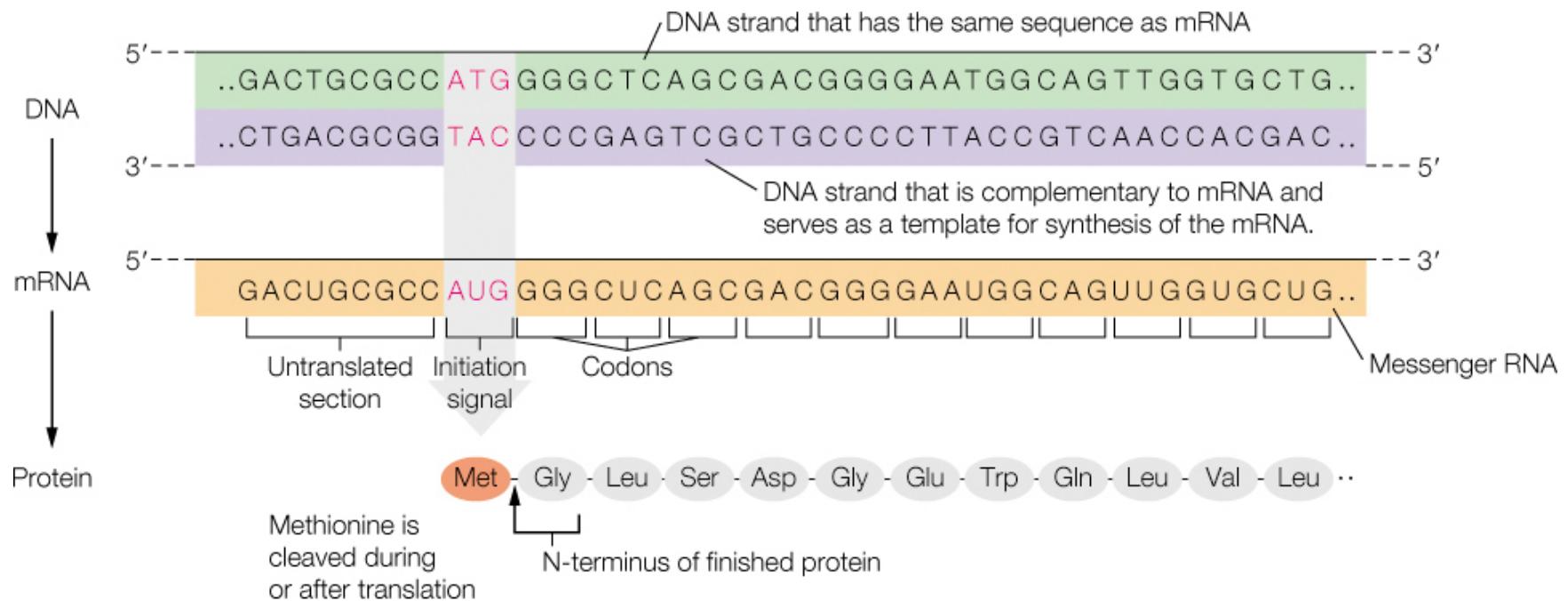


Figure 3-13
© 2013 John Wiley & Sons, Inc. All rights reserved.



Making myoglobin

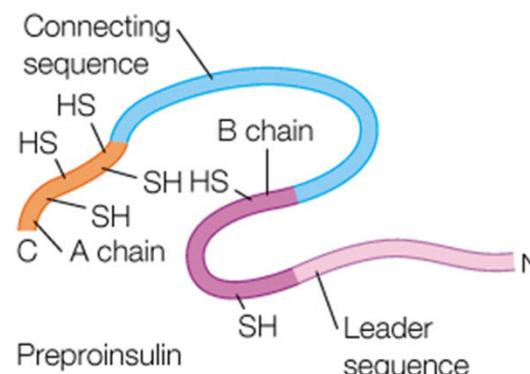
- The myoglobin gene is transcribed into an mRNA which is translated into protein



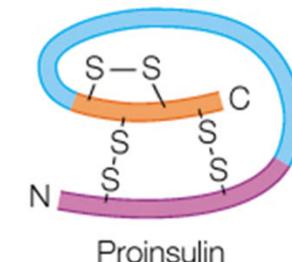
Some proteins have a more complex synthesis

- **Insulin** is synthesized as an inactive precursor, **preproinsulin**, which undergoes enzymatic cleavage to proinsulin, followed by folding and disulfide bond formation, followed by additional enzymatic proteolysis
- This allows insulin, like other peptide or protein hormones, to be synthesized and stored as an **inactive precursor**, available for rapid mobilization on demand

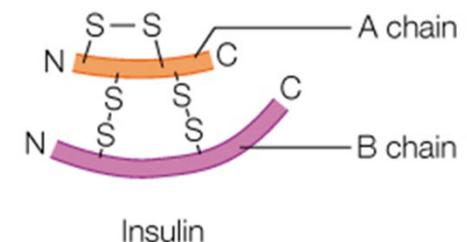
- 1 Preproinsulin is synthesized as a random coil on membrane-associated ribosomes



- 3 Disulfide bonds form



- 4 The connecting sequence is cleaved to form the mature insulin molecule



Amino acids – 2 summary

Stereochemistry and Light absorption

- Amino acids and many other biological compounds are chiral molecules whose configurations are shown by Fischer projections
- The amino acids in proteins all have the L stereochemical configuration
- Proteins absorb UV light at 280 nm

Amino acid derivatives

- The side chains of amino acid residues in proteins may be covalently modified.
- Some amino acids and amino acid derivatives function as hormones and regulatory molecules.
 - Glycine and glutamate are themselves neurotransmitters

Amino acids to proteins

- The peptide bond between amino acids leads to peptides and proteins, with defined sequence and direction
- Proteins vary in size considerably and retain the properties of their constituent side chains.

Amino acids – 2 summary continued

Genes to proteins

- Proteins are encoded by genes: DNA to RNA to protein)
- DNA and RNA are composed of nucleic acids, ribose sugars and phosphate groups
- Nucleotides (base, ribose sugar and phosphate groups) carry energy (e.g. ATP) as well as information.
- DNA and RNA are both made up of **bases** (ATCG or AUCG) supported by a **sugar**, strung together by **phosphodiester** bonds.
 - RNA sugar is **ribose**; DNA sugar is **deoxyribose**.
- Regions of DNA called genes are transcribed into mRNA, which are then translated into proteins.
- After translation, proteins may be modified to generate the biochemically functional form.



Reminders

- Pracs and Tutorials start this week
- Classes 1-3: Pracs this week
 - Pl. do your pre-lab
 - Pl. come to the lab on time, with your lab coat and fully enclosed shoes!
- Classes 4-6: Tutorials this week
 - Class 7: online tutorial
- Please sign up at the text book site from iLearn
 - *Tutorial 1 Quiz (1%) for all classes on the textbook site.*

Proteins 1: Biological Function, Purification and Analysis

Shoba Ranganathan

Applied Biosciences

T: 02 9850 6262; E: shoba.ranganathan@mq.edu.au

Final reminder for Pearson text

- New PDF guide iLearn (also in Discussion Forum).
- Tutorial 1 Quiz (1%) on Pearson only
- All Tutorial Quizzes and Tests will be on Pearson only.
- All Spot Tests from L5 on Pearson only
- No L4 Spot test.

Objectives

Australian Guide to Healthy Eating

- Properties of proteins
 - What do proteins do in cells?
- Protein separation and analysis approaches

AAM: Chapter 5



Proteins: agents of biological function

- Almost every cellular activity depends on proteins
 - “**Proteome**” – studying the entire protein component
 - **Enzymes**: largest group (20%) – catalysts for specific biological reaction(s): upto 10^{16} of the uncatalyzed rate
 - **Regulatory proteins**: 14% - regulate the ability of other proteins to carry out physiological functions
 - ❖ **Hormones**: insulin, somatotropin (pituitary), thyrotropin
 - ❖ **Gene regulatory proteins**: DNA-binding proteins for transcription control
 - **Transport proteins**: 5% - trucking specific cargo from one location to another
 - ❖ **Hemoglobin** for oxygen transport, serum albumin for carrying fatty acids from adipose tissues to organs
 - ❖ **Membrane transport proteins**: facilitate metabolite transport across the membrane *via* mainly channels or pores.

Proteins: biological function - 2

- **Storage proteins:** 5% - reservoir of essential nutrients
 - ❖ **Albumin** (egg white) and **casein** (milk): N storage
 - ❖ **Zeins** (in corn) and **phaseolin** (in peas): N storage
 - ❖ **Ferritin** stores iron in animals
- **Movement proteins:** support cell division, muscle contraction and cell motility
 - ❖ **Actin and myosin:** filamentous proteins for muscle contraction
 - ❖ **Tubulin:** in microtubules (in cell division and in flagella/cilia)
 - ❖ **Dynein and kinesin:** motor proteins for movement of vesicles, granules and organelles along microtubules
- **Structural proteins:** maintain integrity of biological structures as fibres
 - ❖ **α -keratins:** hair, horn, nail
 - ❖ **Collagen:** bone, connective tissue, tendons, cartilage, hide; protective barrier in the extracellular matrix with proteoglycans
 - ❖ **Elastin:** ligaments

Proteins: biological function - 3

- **Signalling proteins:** cellular response to hormones and growth factors in complex networks: signalling pathways
 - ❖ **Hormone receptors** and **protein kinases** which add phosphate groups to other proteins
 - ❖ **Scaffold or adapter proteins:** contains specific modules for recognition and binding of specific structural elements
 - ❖ **Anchoring or targetting proteins:** bind other proteins, causing them to be localized to specific cellular structures
- **Other proteins:**
 - ❖ **Protective:**
 - *immunoglobulins, antibodies* to recognize “foreign” molecules
 - *thrombin, fibrinogen* for clotting blood after injury
 - *Antifreeze proteins*
 - ❖ **Defensive/Exploitive:** toxins and venoms
 - ❖ **Exotic:** Monellin, from an African plant is x2000 sweeter than sugar

Protein Purification & Analysis

Overview

- Environmental conditions such as pH and temperature affect a protein's stability during purification.
- An assay based on a protein's chemical or binding properties may be used to quantify a protein during purification.
- Fractionation procedures take advantage of a protein's unique structure and chemistry in order to separate it from other molecules.

Key Concepts

- A protein's ionic charge, polarity, size, and ligand-binding ability influence its chromatographic behaviour.
- Gel electrophoresis and its variations can separate proteins according to charge, size, and isoelectric point.
- The overall size and shape of macromolecules and larger assemblies can be assessed through ultracentrifugation.

Protein purification needs a strategy

- The cell contains many proteins, in organelles but also in solution and in membranes
- Need to get proteins and other biomolecules out of the cell and into solution
- The protein of interest must be kept as close to its biological environment as possible to retain function.

Factors that affect a protein's stability and function

- **pH:** buffering is usually required.
- **Temperature:** most proteins denature (i.e. become addled) at high temperatures: even a few degrees is enough! Usually 0°C chosen!
- **Degradative enzymes:** proteases and nucleases are released by destroying tissues: temperature/pH control or inhibitors required.
- **Adsorption to surfaces:** many proteins denature at the air-water interface or when in contact with glass/plastic.
- **Storage:** after purification, under inert gas and/or frozen.

Protein purification methods

Methods are based on physical properties:

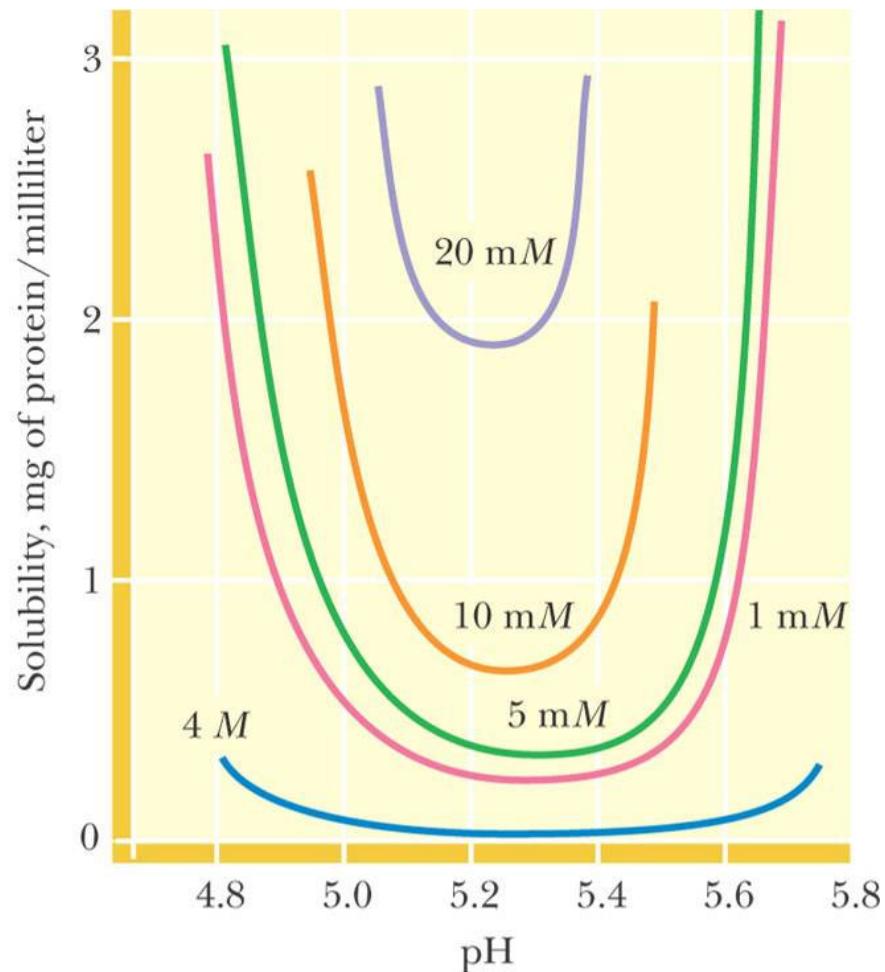
- Solubility of proteins
- Ionic Charge
- Mass
- Size
- Affinity

TABLE 5-2 Protein Purification Procedures

| Protein Characteristic | Purification Procedure |
|-------------------------------|---|
| Solubility | Salting out |
| Ionic Charge | Ion exchange chromatography Electrophoresis Isoelectric focusing |
| Mass | Ultracentrifugation |
| Size | Gel filtration chromatography SDS-PAGE |
| Binding Specificity | Affinity chromatography |

Protein solubility

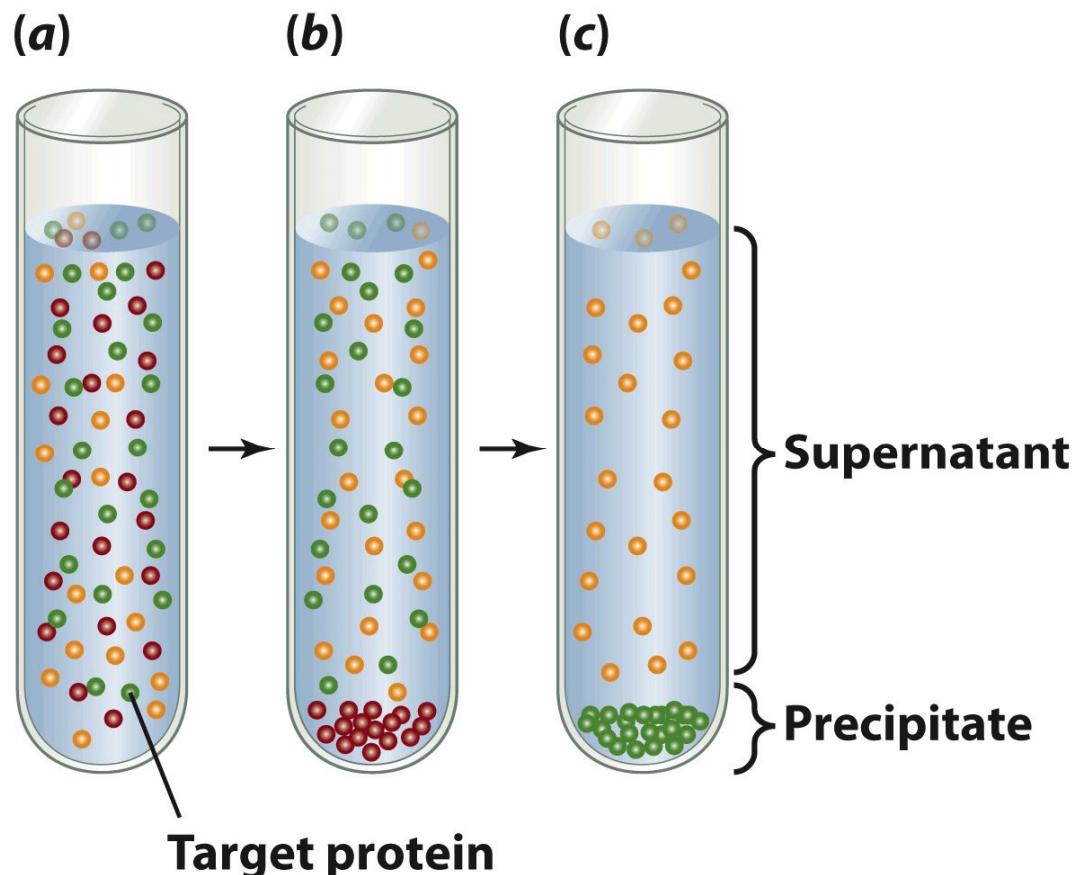
- Proteins are **least soluble at their pI**, when the net charge is zero.
 - At the pI, the electrostatic repulsion between the molecules is a minimum; protein aggregates and precipitates out.
- Generally, solubility increases with ionic strength.
 - The figure shows the solubility of a typical protein as a function of pH and various salt concentrations.



Solubility and ionic strength

- Salting-in
 - Most globular proteins are only slightly soluble in pure water
 - ❖ The **addition of salts** to increase the solubility of these proteins is called **salting-in** (typically ~20 mM)
 - ❖ The added salt induces charge interactions between the protein and the salt, rather than between protein molecules, thus preventing aggregation
- Salting-out
 - **Addition of excessive salt** ($> 1 \text{ M}$) precipitates the protein
 - ❖ The salt removes all the available water and hence the protein precipitates, which is called **salting-out**.
 - ❖ Some salts complex water better than others.
- **Increasing ionic strength at first increases the solubility of proteins (salting-in), then decreases it (salting-out).**
- Also, polyvalent ions act faster than singly charged ions.
 - ❖ Ammonium sulphate is commonly used for solubility-based separation of proteins.
- **Demo in Prac 4**

Fractionation by Salting Out



- Add salt of choice to solution containing the target protein
- Centrifuge and discard unwanted proteins (**red**)
- Add more salt to precipitate the protein of interest (**green**) and centrifuge

Electrophoresis: separation by charge and size

- A sieving method – migration of ions in an electric field.
 - Usually denatures the protein since it is coated with a detergent (sodium dodecylsulfate, SDS), to have a uniform negative charge.
 - Used for proteins as well as DNA.
- Referred to as Polyacrylamide Gel Electrophoresis (PAGE).
 - Separations based on gel filtration (size and shape) as well as electrophoretic mobility (ionic charge).
 - pH ~ 9 so that all proteins are negatively charged and will migrate to the anode, under the applied electric field. Molecules of similar size and charge will move together as a band through the gel.
 - Proteins are visualized on the gel by a stain in the gel, or by X-rays if radioactive, or by antibody binding (Western or immuno- blot).
 - SDS denatures proteins so that separation is by mass alone.

Ion Exchange Chromatography

- DEAE (diethylaminoethyl) cellulose and CM (carboxymethyl) cellulose are widely used resins
- At a pH above the pI or isoelectric point, a protein carries a **negative charge** and will bind to DEAE-cellulose
- Protein can be selectively eluted by applying a gradient of pH or increasing ionic strength

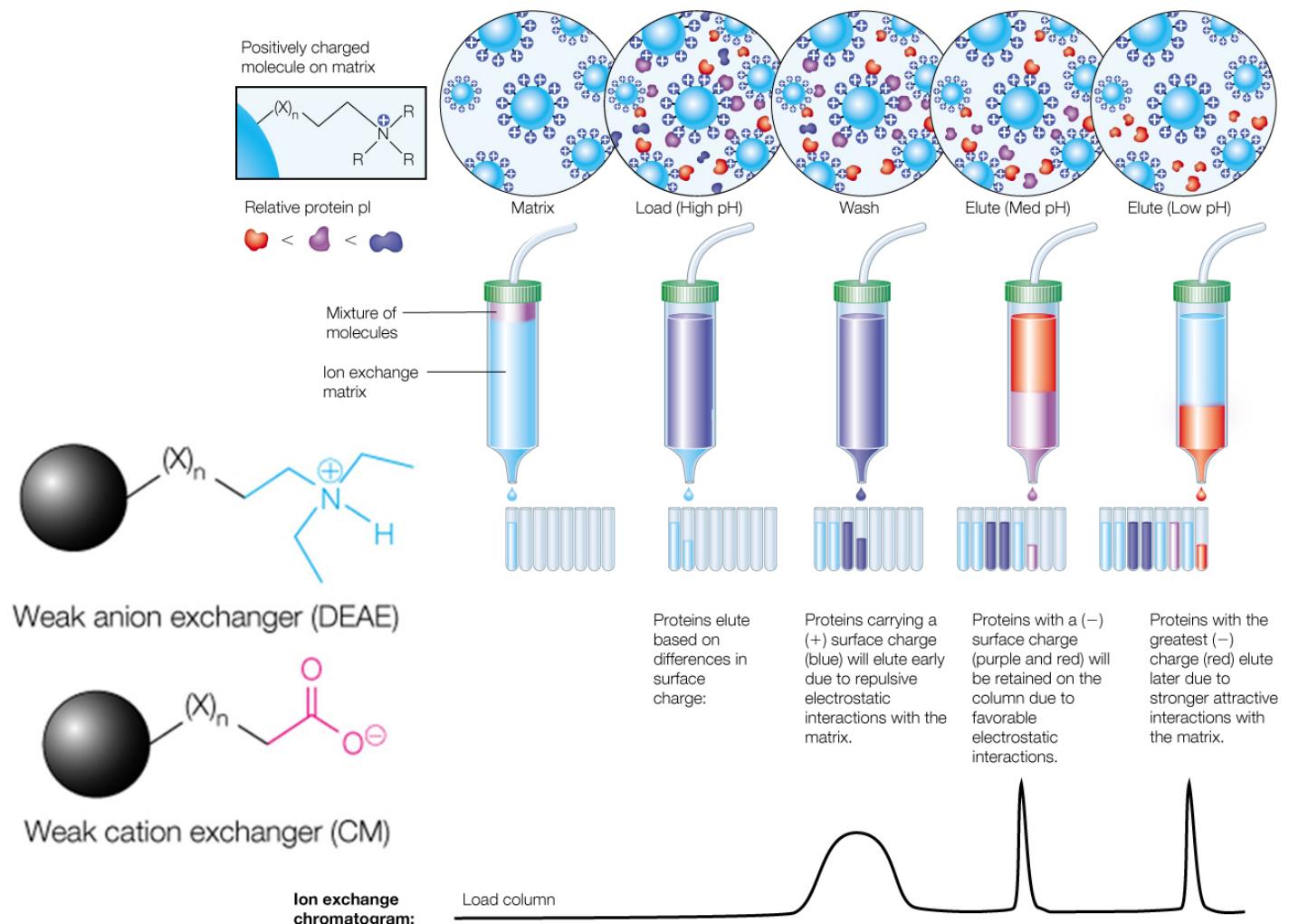


FIGURE 5A.6 An overview of ion exchange chromatography.

Polyacrylamide gel electrophoresis

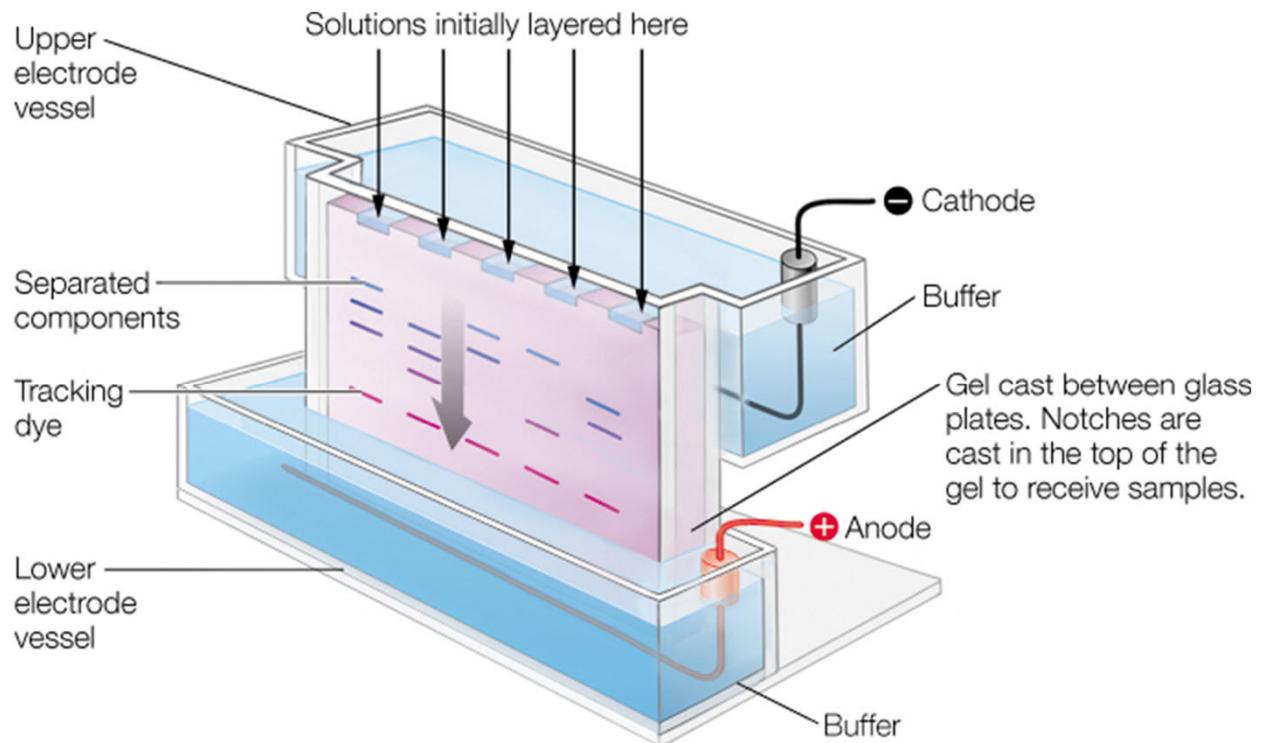


FIGURE 2A.2 Gel electrophoresis.

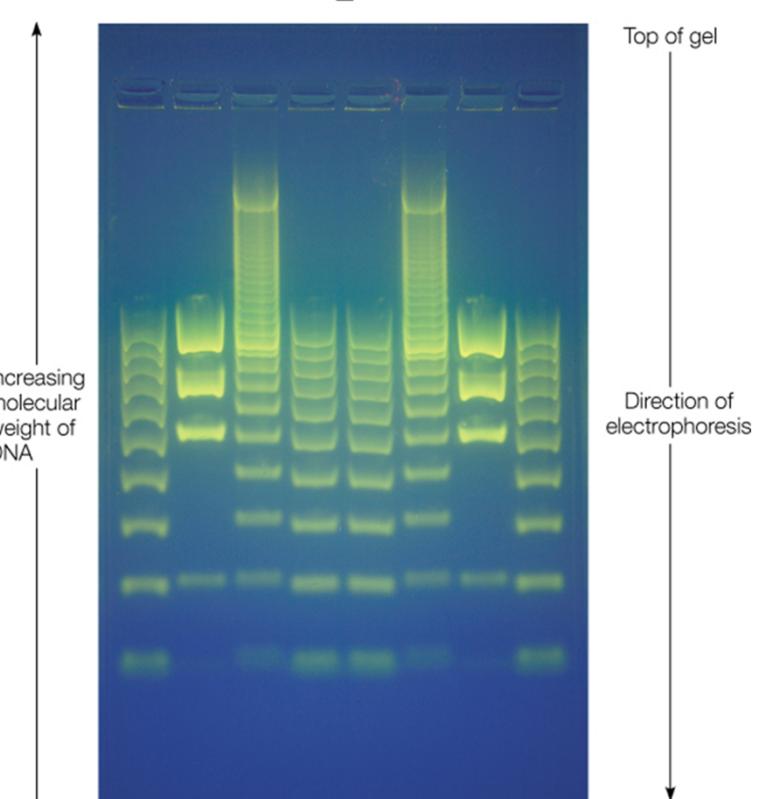


FIGURE 2A.3 + Gel showing separation of DNA fragments.



Isoelectric Focussing

- While agarose and polyacrylamide gel electrophoreses (as used in biochemical laboratories) separate macroions based on their sizes (molecular weights), isoelectric focusing separates macroions based on their surface charge-associated isoelectric point (pl)
 - Macroions move in the electric field until the pH is reached where they have no net charge

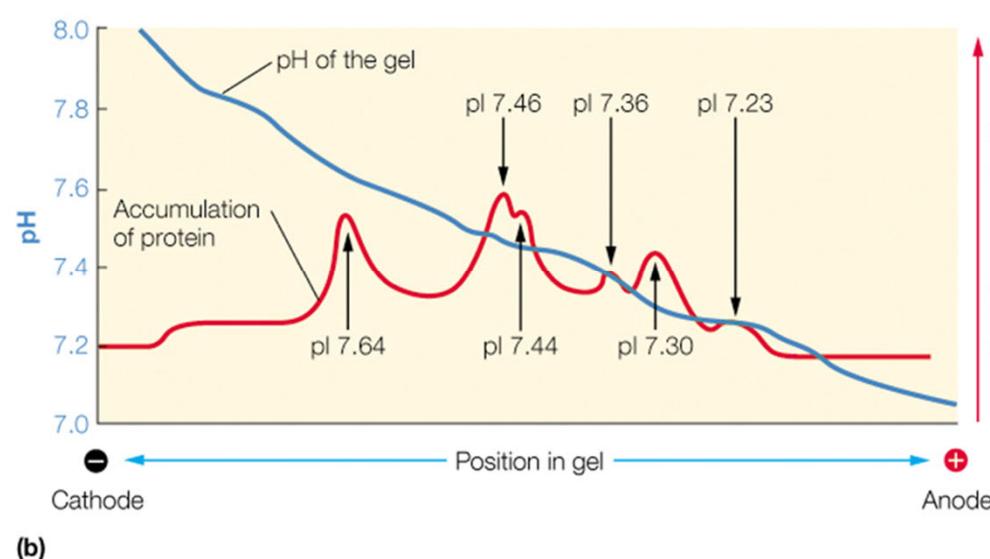
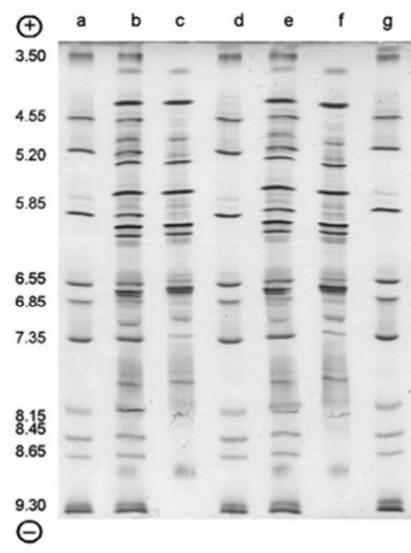
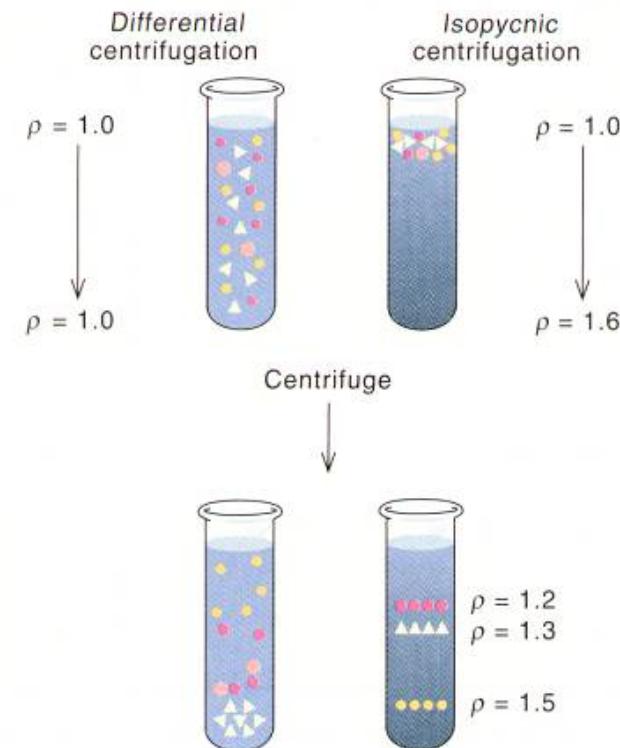


FIGURE 2A.4 Isoelectric focusing of proteins.
(a) pH gradient used from 3.5 (anode end) to 9.3 (cathode end).
(b) Schematic of where proteins with indicated pl values (peaks in red) would appear in the pH gradient (in blue) gel.

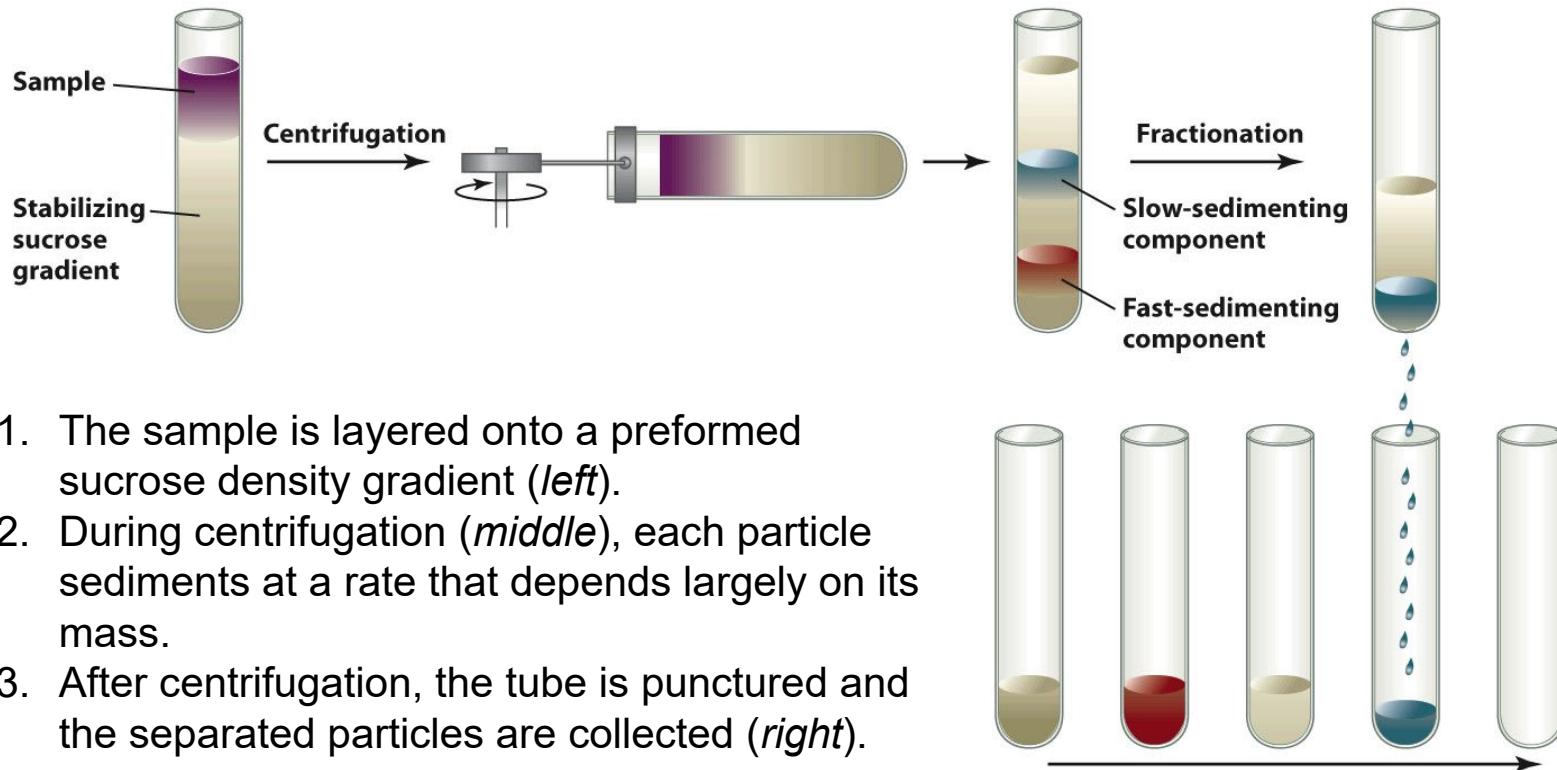


Centrifugation as a separation method

- Differential
 - constant density
 - size dependent
 - time and g force dependent
- Analytical – using an isopycnic solution (with a density gradient)
 - Separation based on density
 - components move to regions of similar density
 - *aka* ultracentrifugation



Ultracentrifugation



1. The sample is layered onto a preformed sucrose density gradient (*left*).
2. During centrifugation (*middle*), each particle sediments at a rate that depends largely on its mass.
3. After centrifugation, the tube is punctured and the separated particles are collected (*right*).

Ultrafiltration membranes are used in dialysis

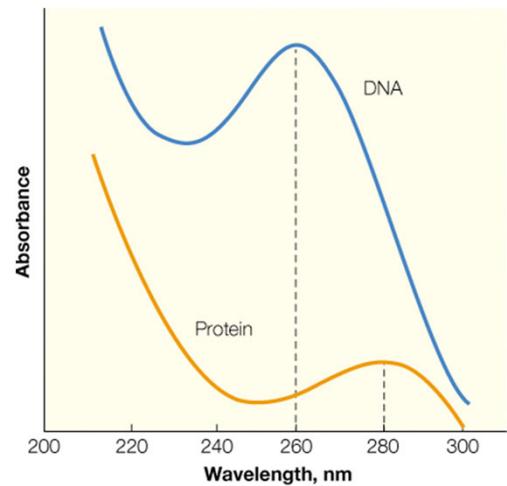
Chromatography: “colour writing” for separation

In general, chromatography is used to separate a mixture of substances, dissolved in a liquid (the “mobile” phase), based on their attraction or affinity for a solid (the “stationary” phase) – originally used to separate coloured substances, hence the name.

1. Paper chromatography uses filter paper as stationary phase and a mixture of solvent and buffer as the mobile phase (**Prac 5 demo**).

2. Column chromatography uses porous substances as stationary phase (called the matrix). There are different column types depending on the nature of substances to be separated.

- Presence of proteins can be monitored by collecting fractions of the mobile phase and determining the protein concentration by **absorption spectrophotometry at 280 nm**.
- Plotting absorbance at 280 nm gives graph (chromatogram) with **peaks for each protein** separated by the column.



Column Chromatography

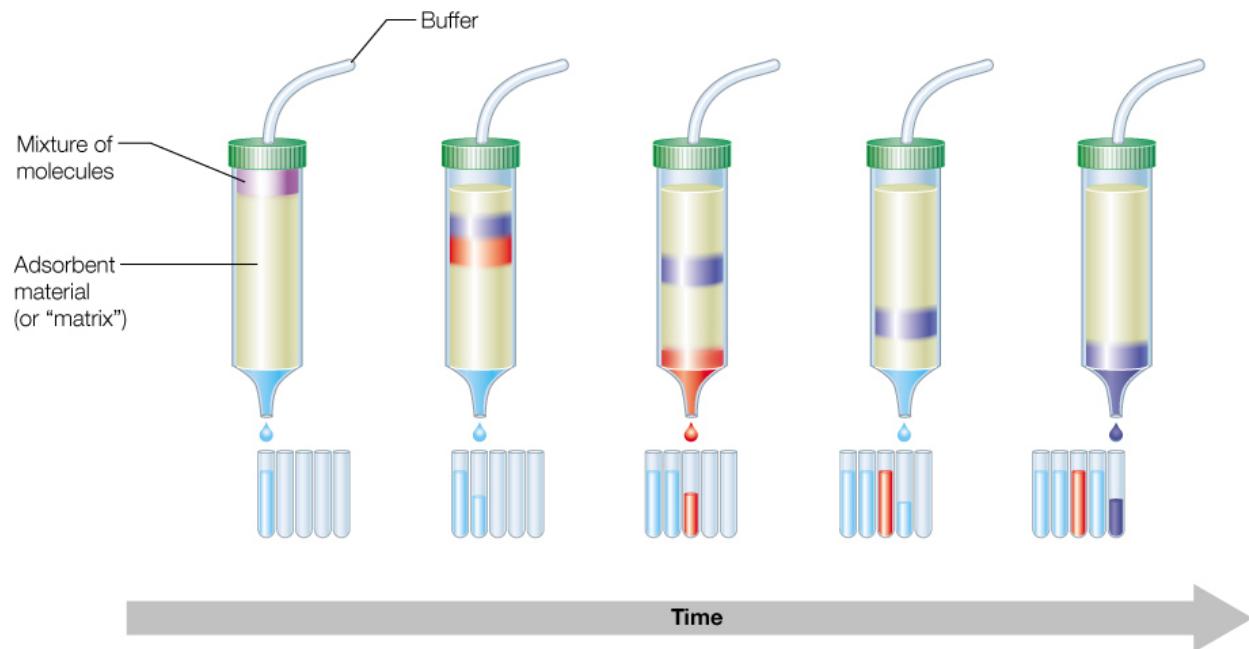


FIGURE 5A.2 The principle of column chromatography.

- A **mixture** of (cellular) proteins is separated as a result of differential interactions with the column **matrix**.
- The **buffer elutes** (or moves) the components of the mixture
- The more a protein interacts with the column material or **matrix**, the later it will elute from the column

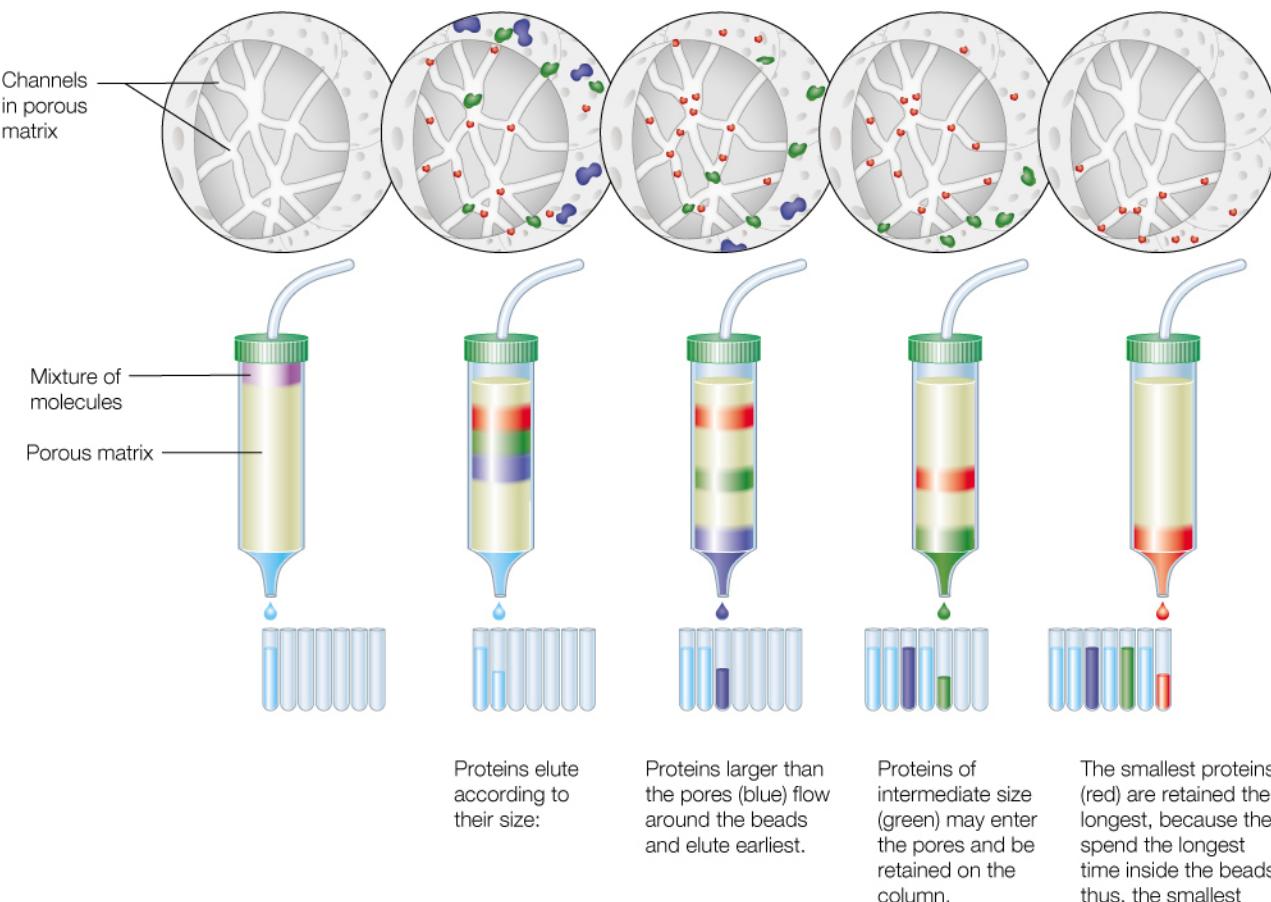


FIGURE 5A.7 An overview of size exclusion chromatography.

Size Exclusion Chromatography

- Separation of proteins is based on the apparent sizes of the molecules
- Larger proteins elute earlier because they are excluded from the interior volume of the chromatography matrix
- Smaller proteins elute later because they are retained within the pores of the chromatography matrix, resulting in a longer residence time on the column
- Prac 4 Theory**

Affinity Chromatography

- Selective adsorption of a protein to a natural or synthetic ligand, usually a substrate or inhibitor
 - The matrix contains the covalently bound ligand
- The protein interacts strongly with the **affinity matrix** and will elute last
- For recombinant proteins, a short His tag of 6 His residues is added in, which will bind preferentially to metal ions in an immobilized metal affinity chromatography (**IMAC**) column.

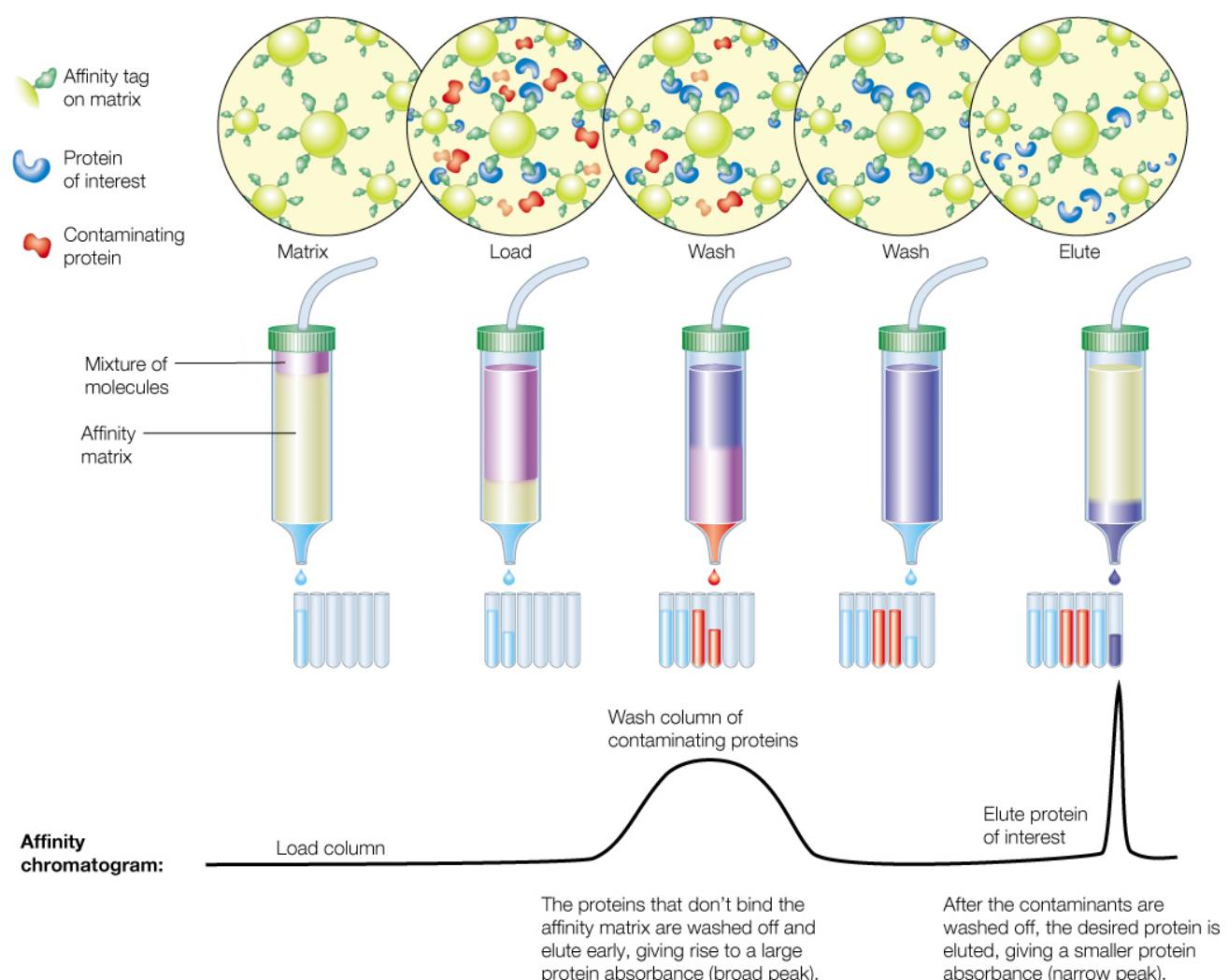


FIGURE 5A.3 An overview of affinity chromatography.



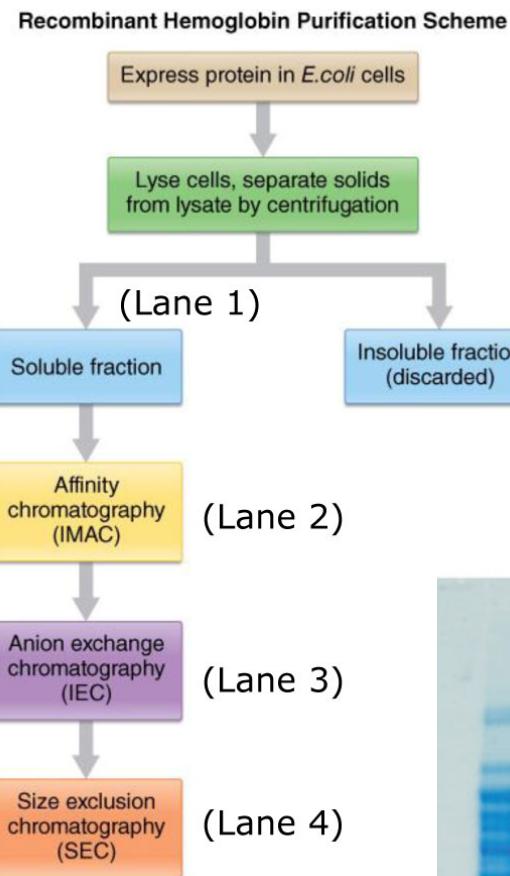


FIGURE 5A.8 Flowchart for the purification of recombinant hemoglobin.



MACQUARIE
University

Example: Purifying a Recombinant Hemoglobin Mutant

Haemoglobin (Hb) has two polypeptide chains: α -globin and β -globin. This mutant has an 8-amino acid insertion in the β -globin sequence. After protein production, the cells are lysed and the contents of the cytoplasm released into a buffered solution.

Centrifugation: removes the insoluble material (e.g., membranes and precipitated protein aggregates) from soluble fraction.

Lane 1: The resulting supernatant is a complex mixture of nucleic acids and proteins.

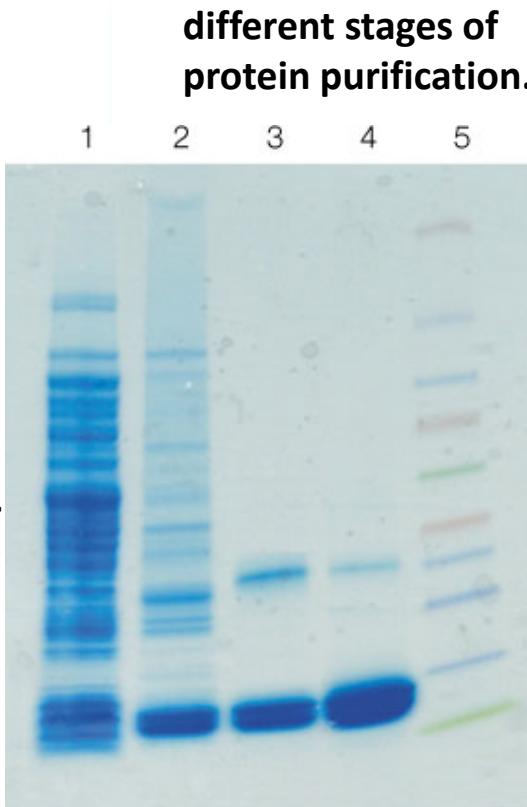
Lane 2: an IMAC purification step significantly purifies the mutant Hb.

Lane 3: several contaminants are removed by anion exchange on Q resin.

Lane 4: Finally, SEC is used to separate the Hb from the contaminant at 35 kDa. Lane 4 is overloaded in order to detect the presence of impurities in the final preparation.

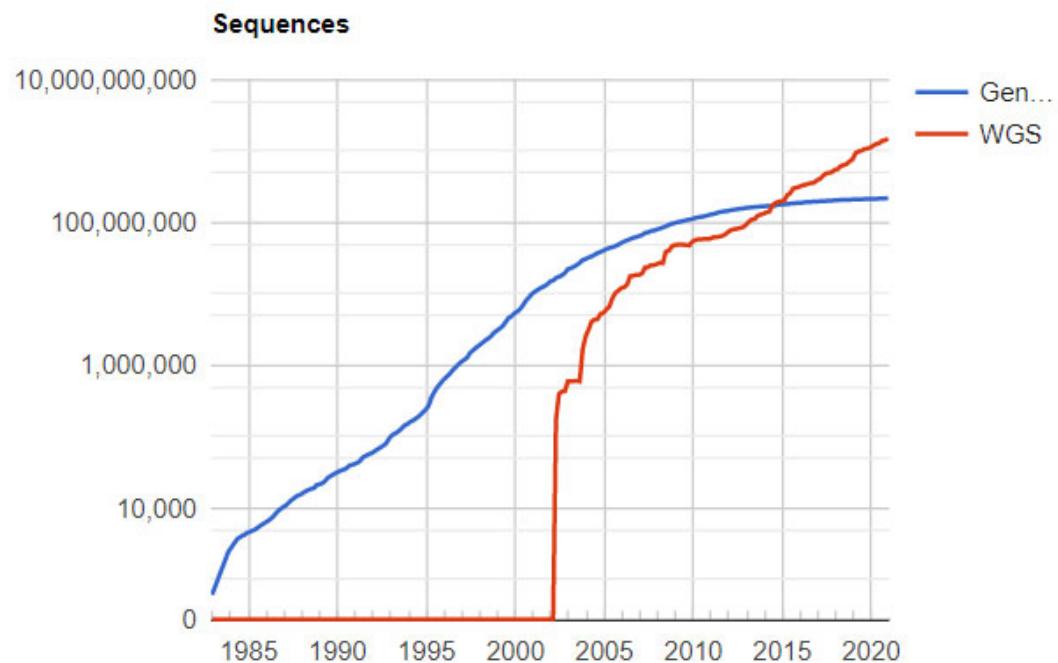
(Lane 5 contains a series of known mass standards)

17 kDa In this case, the hemoglobin mutant is greater than 95%
10 kDa of the total protein in the purified sample.



From Gene Sequence to Protein Function

- Genes code for proteins and
- Late 1970s: recombinant DNA technology became available to express many copies of a single gene which then lead to protein expression.
 - New field of Systems Biology
 - Millions of DNA sequences now deposited in databanks: GenBank now has 221,467,827 sequences
 - Whole genome shotgun (WGS) sequencing has led to 1,517,995,689 sequences! (as of Dec. 2020)



However, Biochemistry is interested in what these sequences do! i.e. their function!

Using sequences to find protein function

- Bioinformatics to the rescue!

- Gene sequences can be translated on a computer and then matched against these millions of database sequences to pick up functional clues
- If a new sequence matches the human myoglobin, it is possibly a globin?
- Let's see how well myoglobin and haemoglobin match up (they both contain ***heme*** – a ligand)

Score = 30.8 bits (68), **Expect = 6e-06**, Method: Compositional matrix adjust.

Identities = 32/133 (25%), Positives = 48/133 (37%), Gaps = 40/133 (30%)

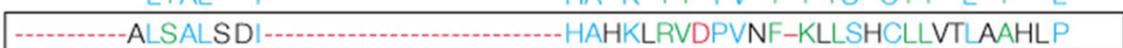
| | | | |
|----------------|-----|--|-----|
| Human Mb | 2 |  | 61 |
| Human α | 2 |  | 46 |
| Human Mb | 62 |  | 120 |
| Human α | 47 |  | 82 |
| Human Mb | 121 |  | 133 |
| Human α | 83 |  | 95 |

FIGURE 5.21 BLAST alignment between human myoglobin and human A globin sequences.



How similar are 2 sequences?

- Identities: 25% i.e. same residues
- Positives: this includes chemically similar side chains and so the score goes up to 37%
- Gaps: 30% - not so good, but maybe the sequences diverged a lot during evolution
- Expect = 6e-06: this is the expectation value for finding this match by chance: probability score and the smaller it is, the better.

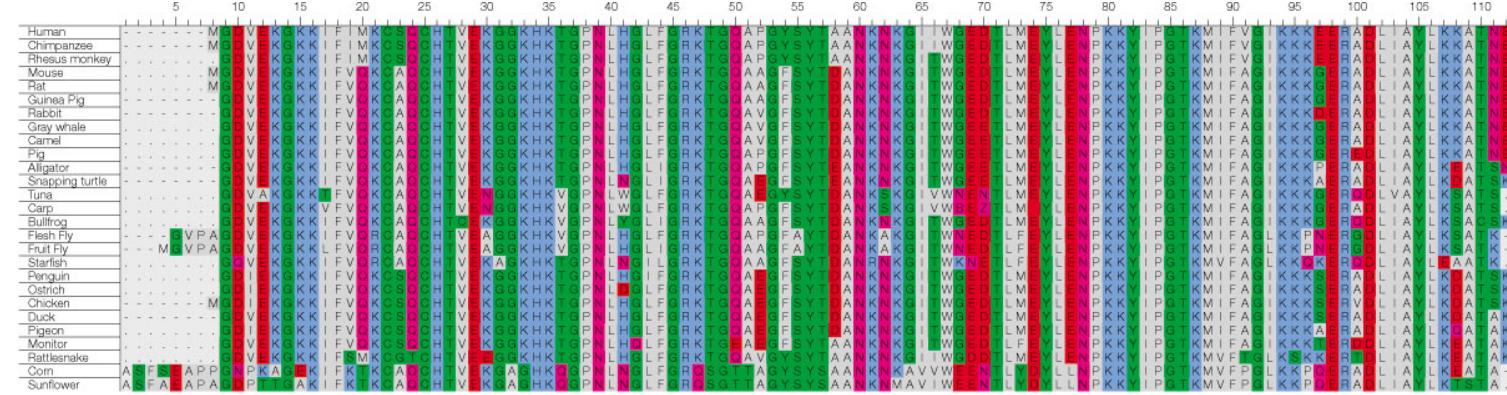
Score = 30.8 bits (68), **Expect = 6e-06**, Method: Compositional matrix adjust.
Identities = 32/133 (25%), Positives = 48/133 (37%), Gaps = 40/133 (30%)

| | | | |
|----------------|-----|---|-----|
| Human Mb | 2 | LSDGEWQLV LNV WGKV EADIPG HGQEVL RLF KGH PET LEK FDK F KHL KSE DEM KASE DL | 61 |
| Human α | 2 | LS PAD KTN V KAA WGKV GA HAGE YGA EA LER MF LS F PTT KTY FPHF ----- | 46 |
| Human Mb | 62 | KKHGATV LT ALGG I LKKKGH HEAEIK PLA QSH A TKHKI - PV KYLE FISE CI IQ VLQ SKHP | 120 |
| Human α | 47 | L + AL I HA K ++ PV + + + S C ++ L + L ----- ALSA L SDI ----- HA HKL RV DP VN F KLL SH C LL VT LA AH LP | 82 |
| Human Mb | 121 | GDF GADA QG AM NK + F + + + K | 133 |
| Human α | 83 | A EFT PAV H AS LDK | 95 |

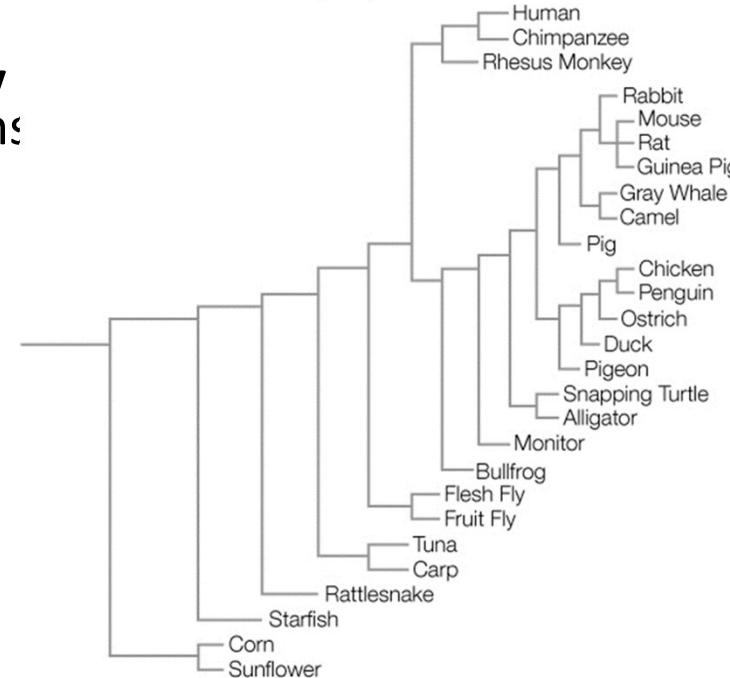
1. So 25% is good and if two sequences are 25% identical, they are considered homologous.
 - Likely to have the same structure, i.e. folded shape and
 - The same function!
2. The probability score helps us to find homologous sequences (< 1e-05).

Homology search

- Matching to an entire database finds homologous sequences from different organisms
- Sequence alignment of cytochrome c, a **mitochondrial respiratory protein**, from 27 organisms
- The more similar two homologous protein sequences are, the more closely they are related evolutionarily
- Lining them up (i.e. their alignment) can help us trace evolutionary history of a protein/gene.



(a) Alignment of cytochrome c sequences from 27 organisms, where hydrophobic amino acids are highlighted in gray, basic amino acids in blue, acidic amino acids in red, and polar uncharged amino acids in green (except for Asn and Gln, which are magenta).



(b) A phylogenetic tree for the sequences shown in (a). Branches indicate points of evolutionary divergence based on differences in the amino acid sequences of the aligned proteins.

FIGURE 5.22 Sequence alignment and a phylogenetic tree for cytochromes c from different organisms.

- Mitochondrial respiration uses homologous proteins in multiple organisms!
- Sequences that are very similar are from closely related organisms – lots of use in biochemistry, biotechnology and of course, biology!
- The function of a protein is conserved even though they look slightly different!

Consensus Sequences from an alignment

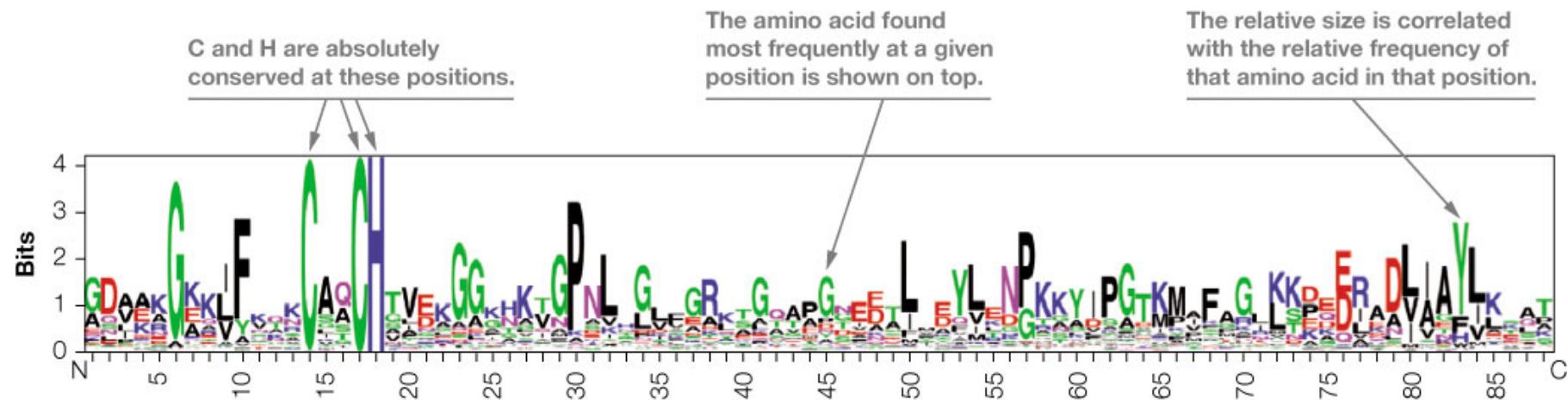
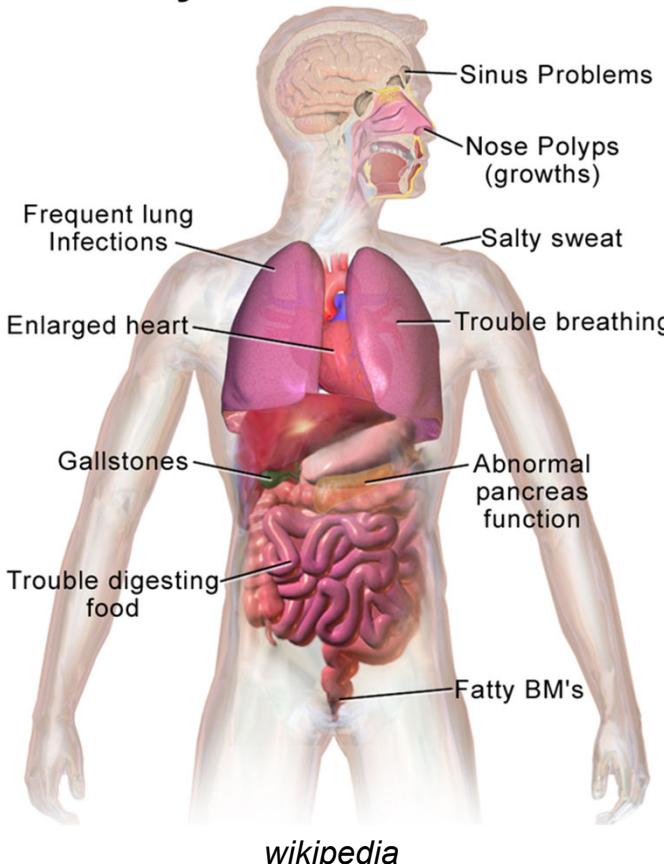


FIGURE 5.23 Sequence logo from the alignment of 412 sequences from the cytochrome c family.

- A consensus sequence (here for cytochrome c) allows us to identify the most highly conserved amino acid residues – these may be in separate clusters!
- These are critical for function and mutating them might completely destroy biological activity
- Naturally occurring mutations or variations have led to the discovery of genetic diseases

Example: Cystic Fibrosis (CF)

Health Problems with Cystic Fibrosis



- Genetic disorder that affects mostly the lungs but also the pancreas, liver, kidneys and intestine. Long-term issues include difficulty in breathing and coughing up sputum as a result of frequent lung infections.
- Recognized as a specific disease by Dorothy Andersen in 1938, with descriptions that fit the condition occurring at least as far back as 1595!
- CF is caused by a mutation in the gene cystic fibrosis transmembrane conductance regulator (CFTR). The most common mutation, **ΔF508**, is a deletion (Δ signifying deletion) of three nucleotides that results in a loss of the amino acid phenylalanine (F) at the 508th position on the protein.
- $\Delta F508$ -CFTR, which occurs in >75% of patients globally, creates a protein that does not fold normally and is degraded by the cell.



Summary

- Proteins are polymers of α -amino acids
 - twenty common amino acids (and two rare), distinguished by their side chains (R groups), are incorporated into proteins
- Proteins are produced by condensation of amino acids via peptide bond formation
- The unique defined sequence of amino acids constitutes the primary structure of proteins
- In cells, genes are transcribed into messenger RNA (mRNA), which is then translated into a polypeptide strand at the ribosomes

Summary

- Using recombinant expression techniques, proteins can be produced in non-native organisms at very high yields
 - Genes of interest are incorporated into these non-native organisms, and then transcribed and translated by these organisms
- Depending on the certain properties of the produced proteins, they can be purified at high purity by various chromatographic techniques, and the protein masses and sequences can be determined.

Reminders

- Pracs and Tutorials start this week
- Classes 1-3: Pracs this week
 - Pl. do your pre-lab
 - Pl. come to the lab on time, with your lab coat and fully enclosed shoes!
- Classes 4-6: Tutorials this week
 - Class 7: online tutorial
- Please sign up at the text book site from iLearn
 - *Tutorial 1 Quiz for all classes on the textbook site.*

Proteins 2: Sequencing and Secondary Structure

Shoba Ranganathan

Applied Biosciences

T: 02 9850 6262; E: shoba.ranganathan@mq.edu.au

Objectives

- How to determine the order of amino acids in a protein (i.e. sequence or primary structure)
- Where is protein sequence information deposited and shared?
- What are domains?
- Common local (secondary) structures in proteins.

Protein Sequencing Workflow

Current sequencing combines:

1. Sanger's sequencing approach for bovine insulin (1st Nobel Prize) and
2. Mass spectrometry for finding the masses of fragment peptides
 - Masses are then identified with specific sequences.

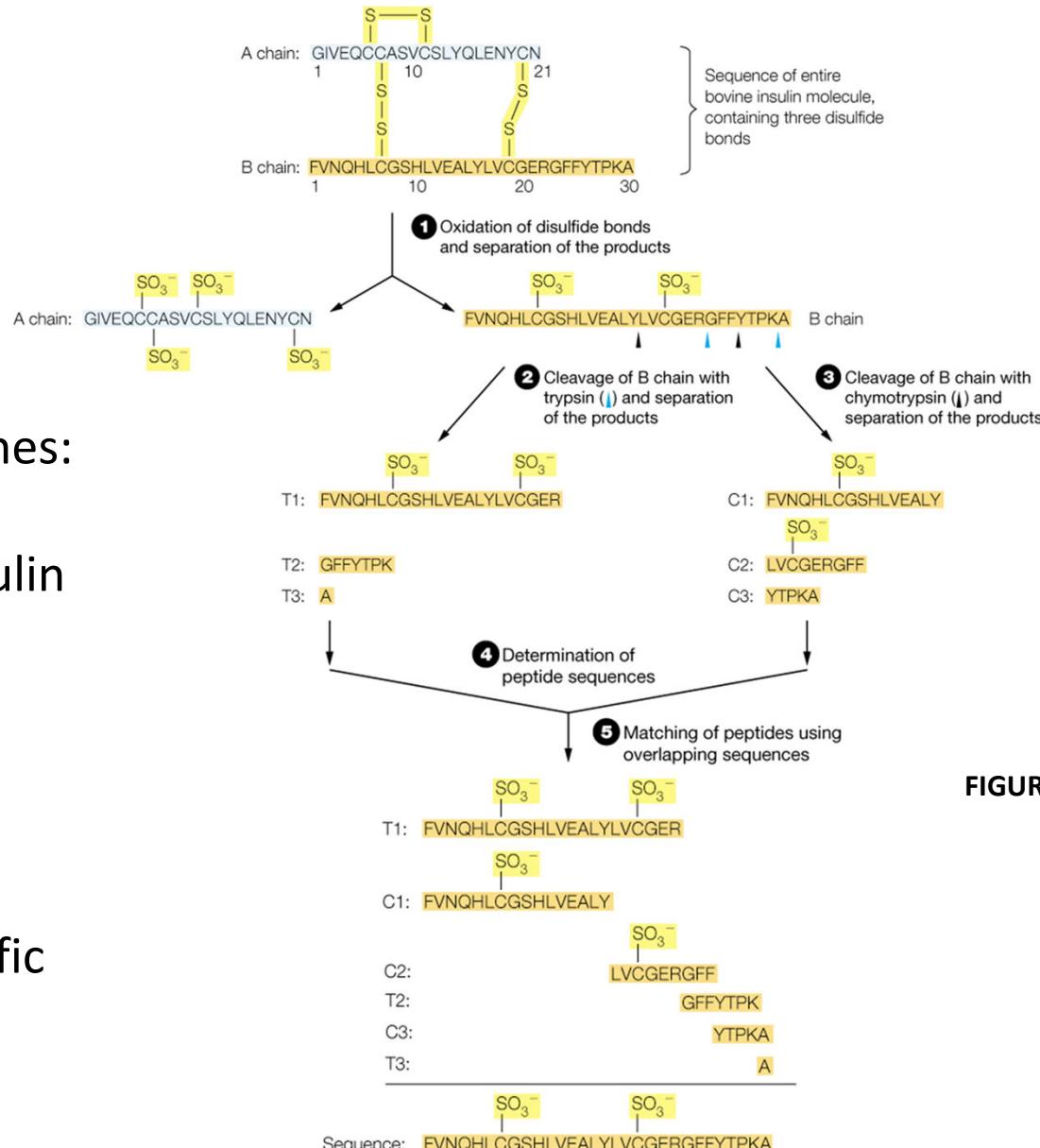
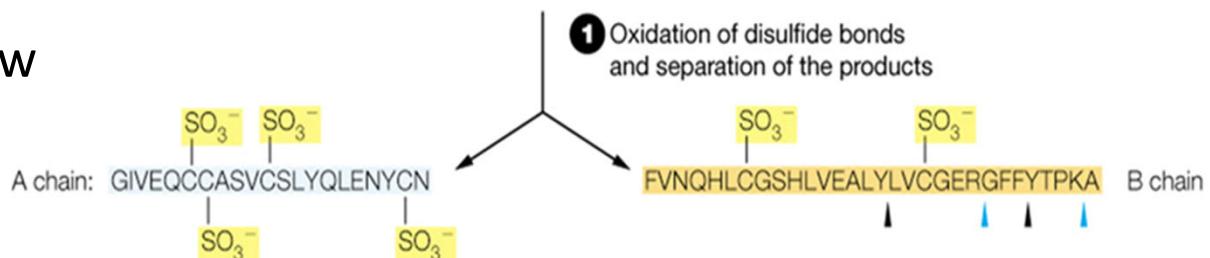
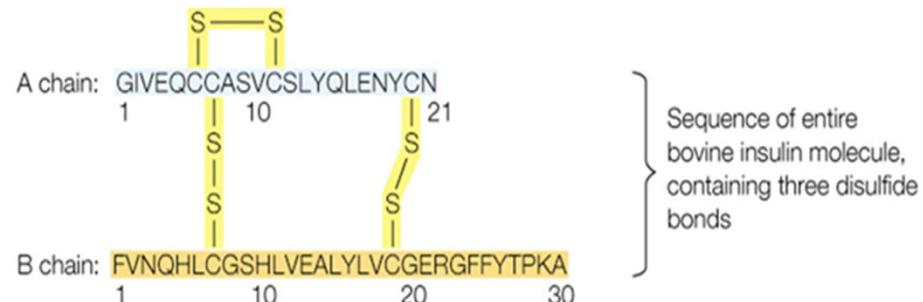
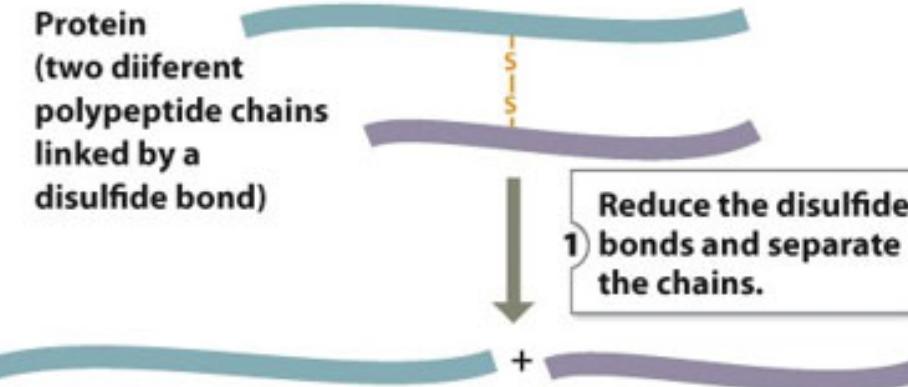


FIGURE 5B.4 Sequencing the

Step 1!

Insulin is made up of two protein chains held together by disulphide bonds

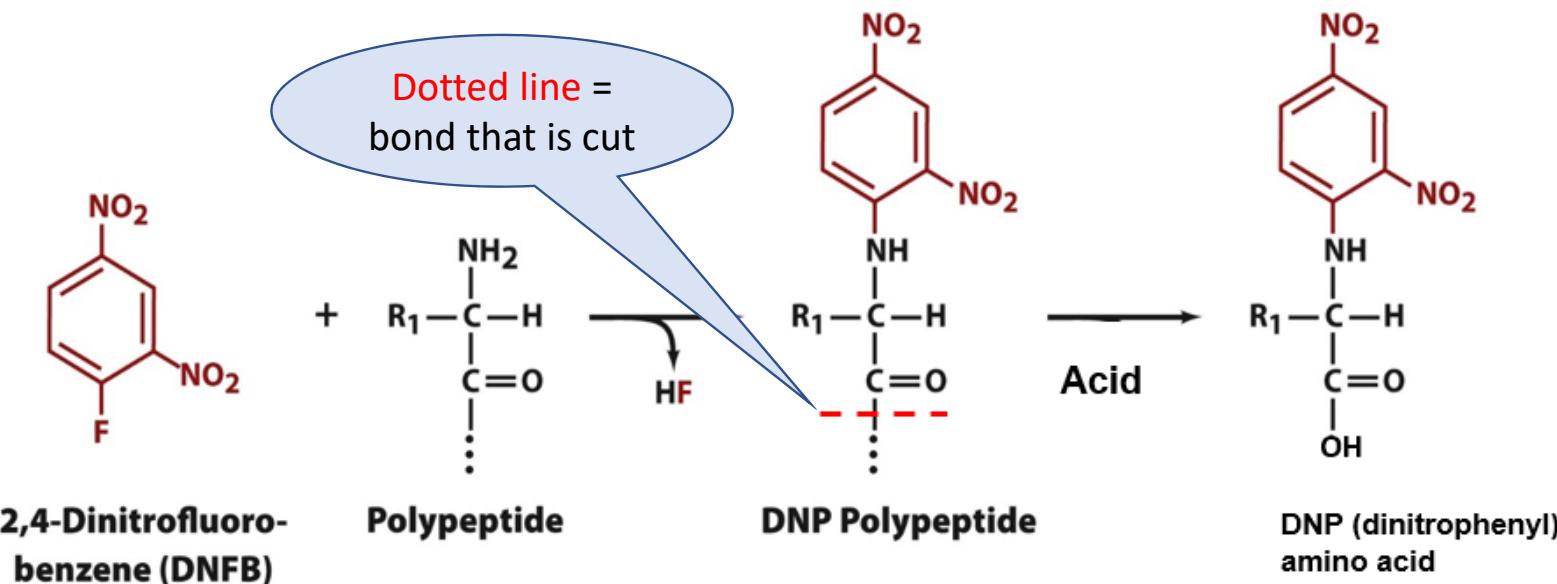
1. Locate the disulphide bonds and
2. Separate the chains by oxidising the disulphide bonds.
3. But we need to know how many chains there are!



How many protein chains are there?

1a. Amino terminal identification:

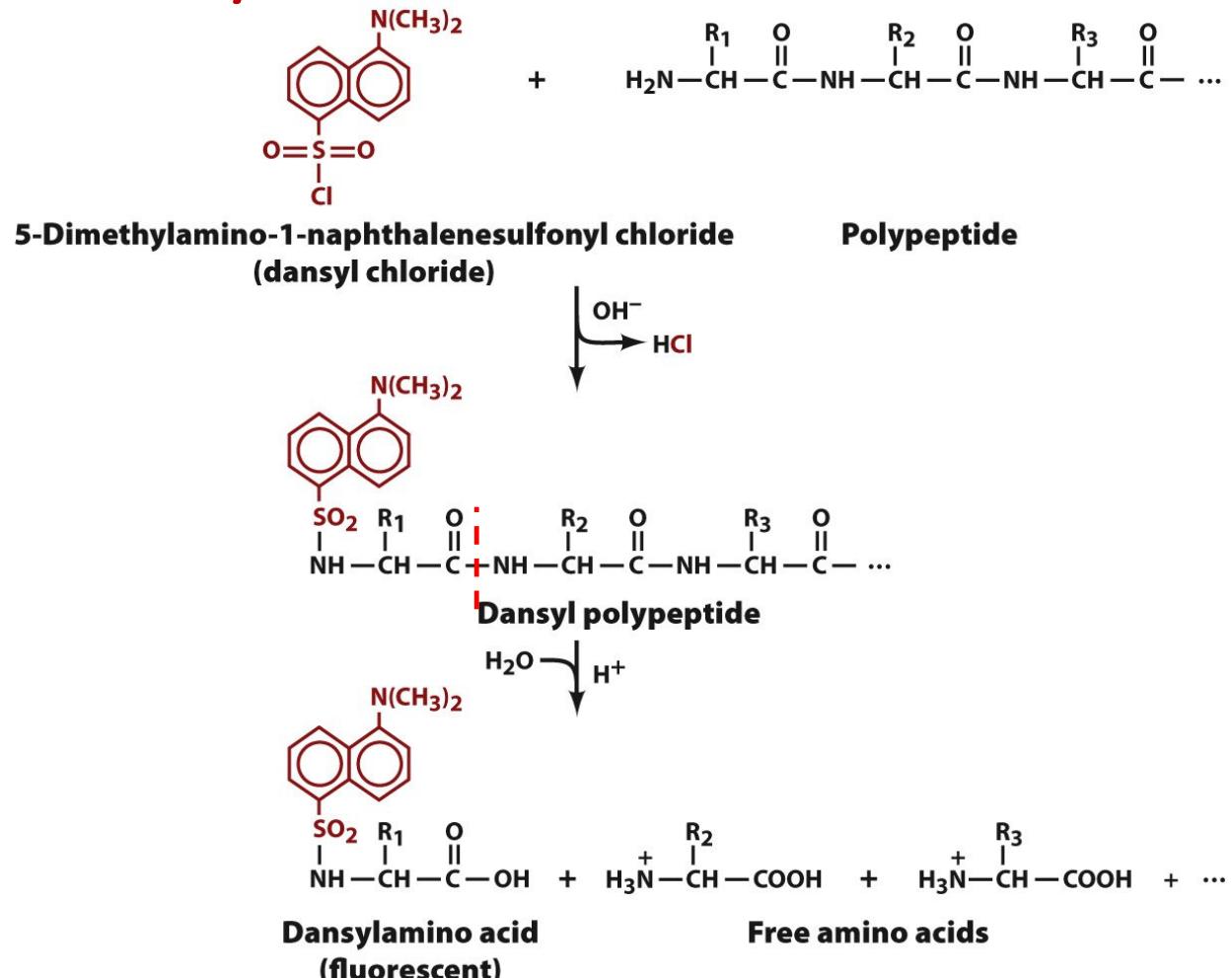
Dinitrofluorobenzene or Sanger method



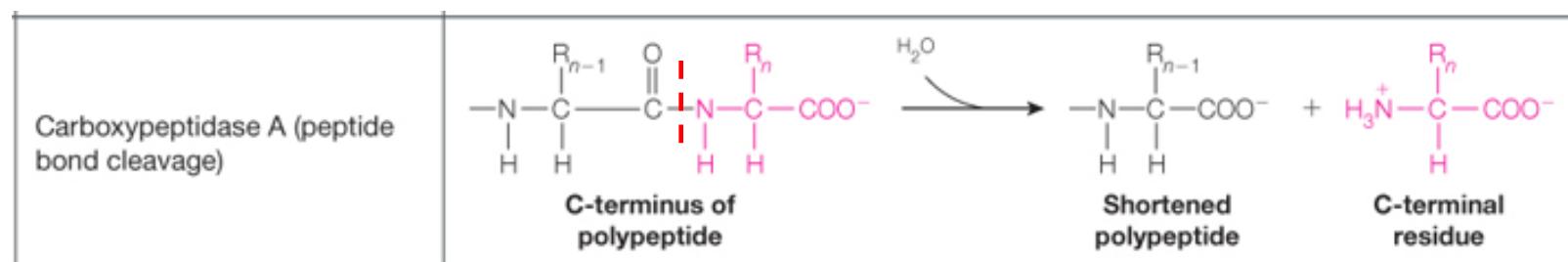
- Took 10 years to sequence insulin - Nobel Prize in 1958
- Currently sequencers take much less time
- Similar method used for sequencing DNA - second Nobel Prize in 1980



1b. Amino terminal identification: Dansyl Chloride Method



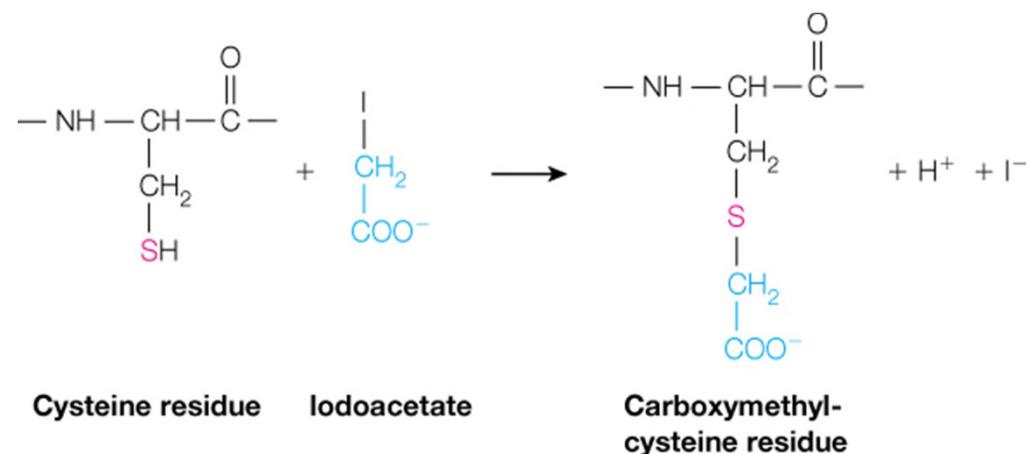
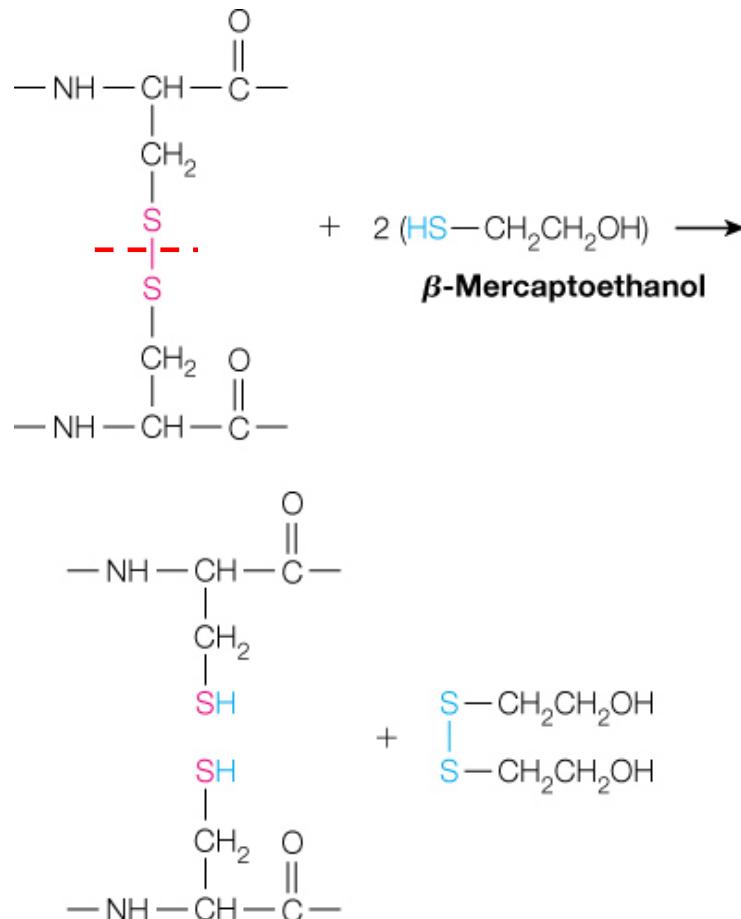
2. Carboxyl terminal identification: Carboxypeptidase



Extracted from Table 8.1

Once we know how many different N- and C-terminal residues there, we can work out if we have more than one protein chain.

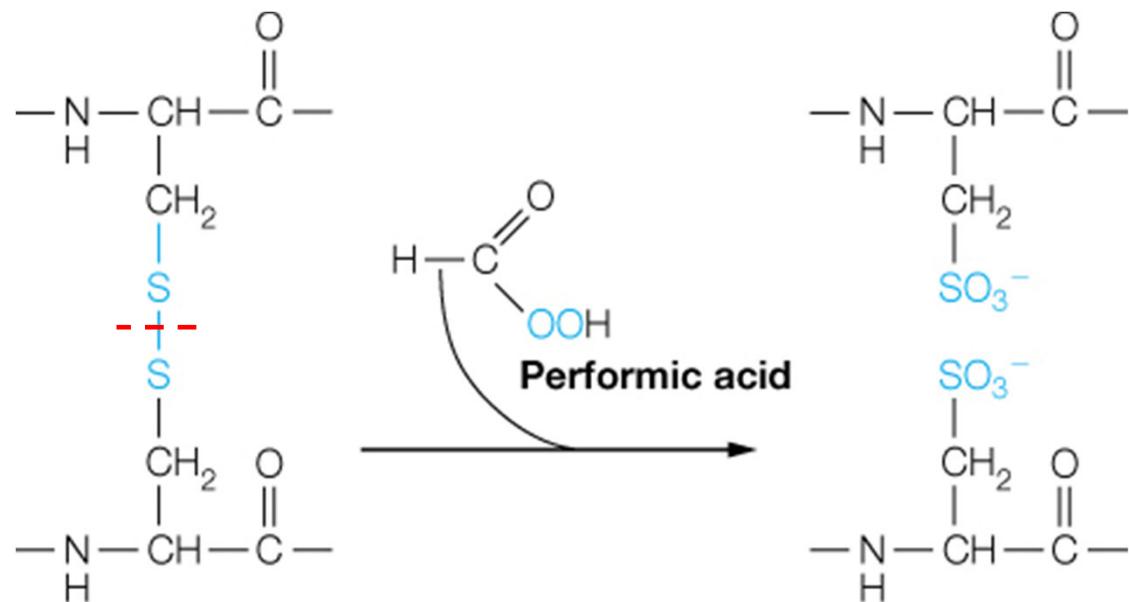
3a. Breaking Disulfide Bonds – 2 steps



Chapter 5, Unnumbered Figures 2 & 3, Page 173

1. Reversible and mild disulphide bond cleavage by **β -mercaptoethanol**
2. The free –SH ends are protected (or blocked) by **iodoacetate** treatment to prevent the free cysteine side chains from forming the same or other disulphide bonds.

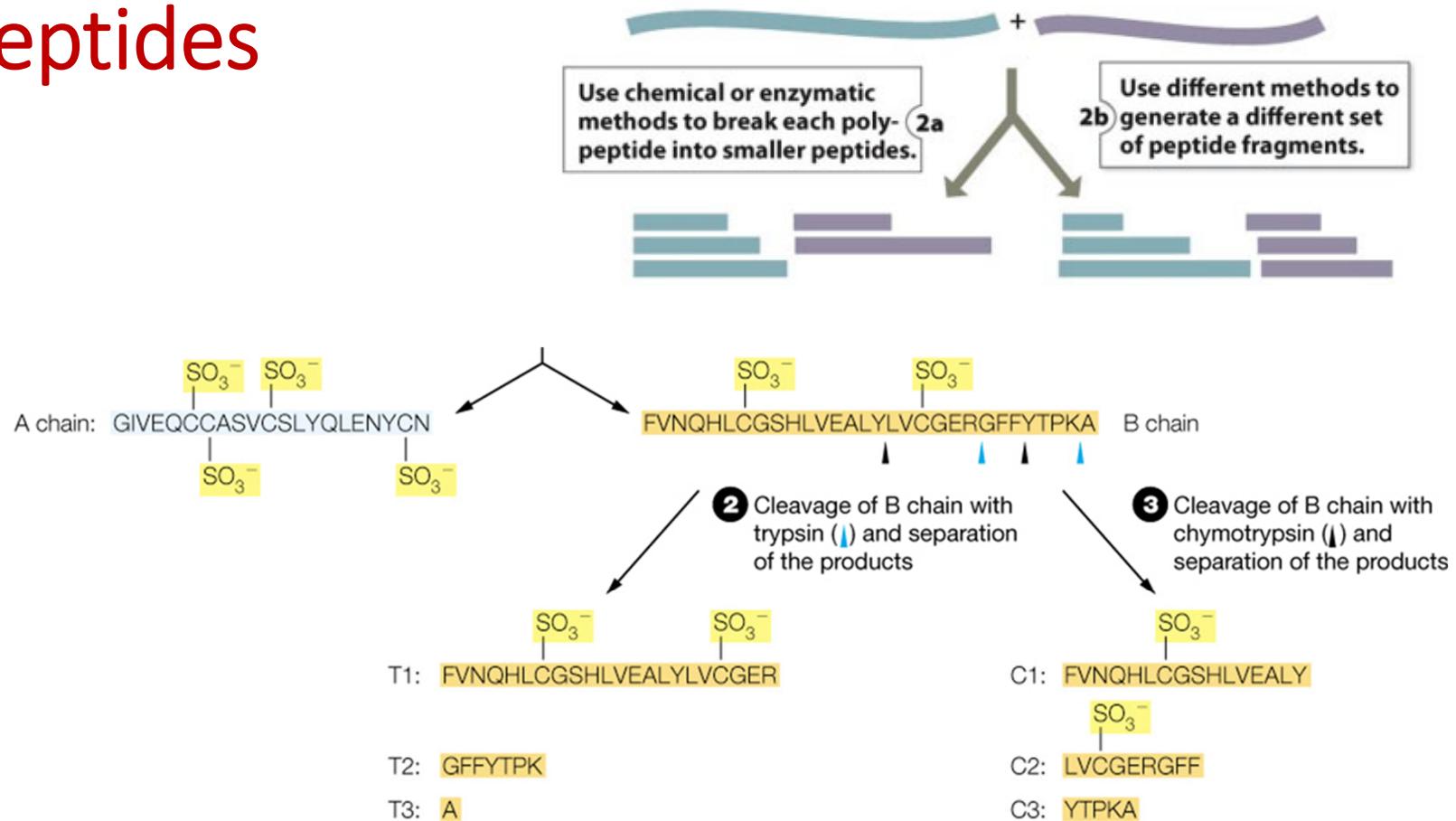
3b. Breaking Disulfide Bonds – 1 step



Chapter 5, Unnumbered Figure 1, Page 173

Performic acid provides irreversible oxidation of disulphide bonds.

Step 2: cutting up the separated protein chains to peptides



Proteins to Peptides - 1: Cyanogen bromide (Met-specific)

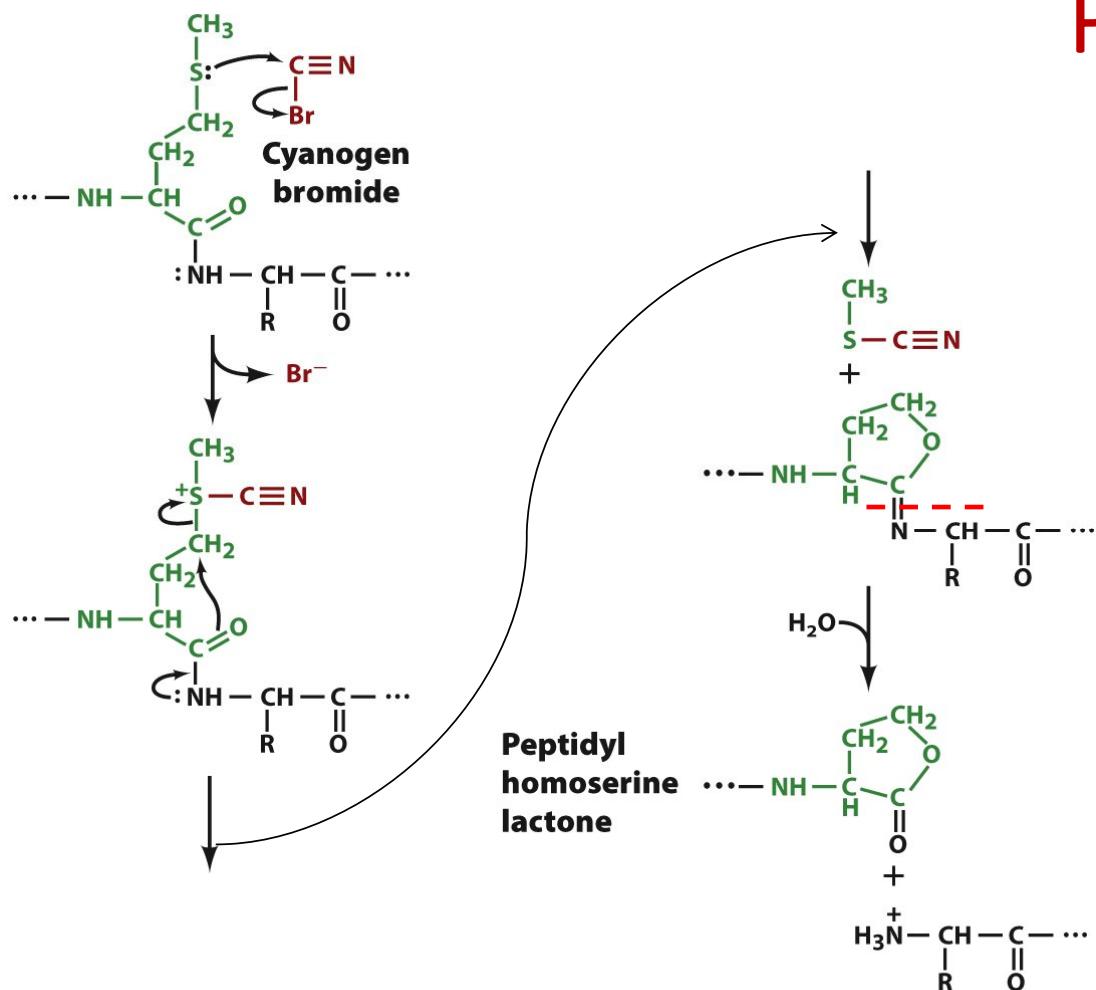
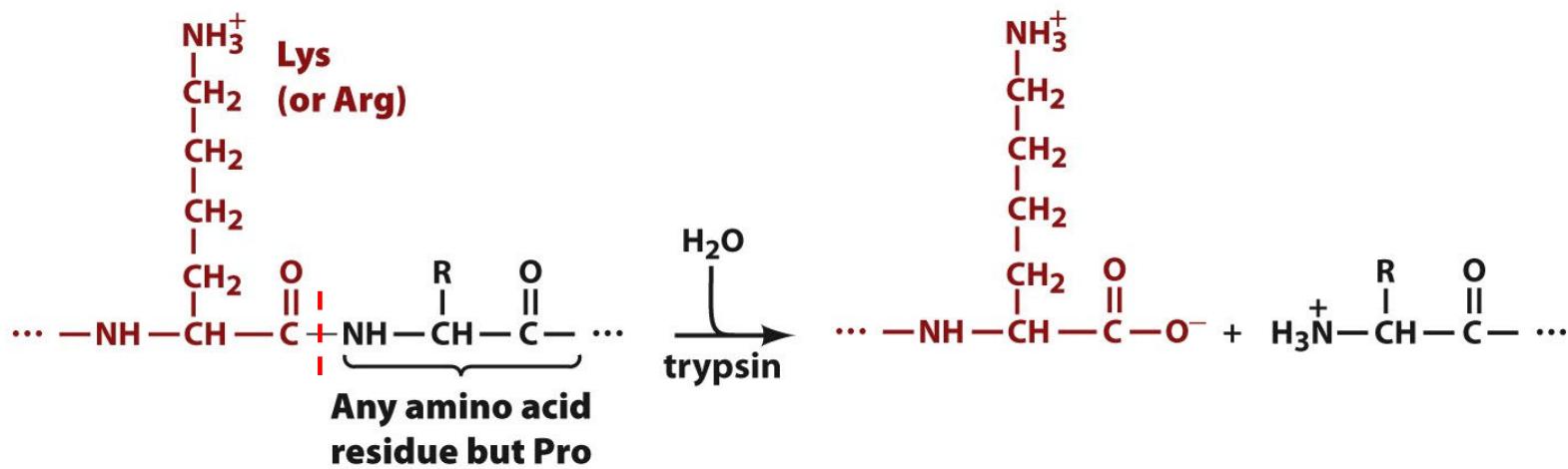


Figure 5-15
© 2013 John Wiley & Sons, Inc. All rights reserved.



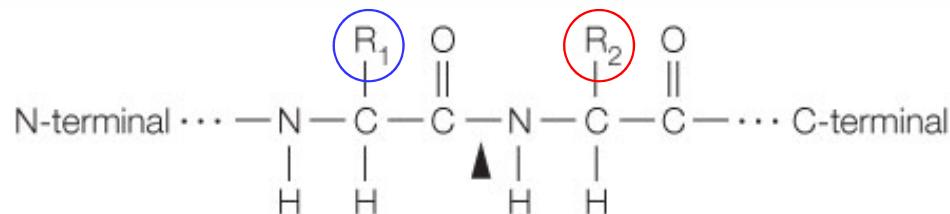
Proteins to Peptides - 2: Specific endopeptidases



Most important method to **fragment proteins**: used in proteomics

TABLE 5.3 The sequence specificities of some proteolytic enzymes and cyanogen bromide

Breaking peptide bonds



| Enzyme | Preferred Site ^a | Source |
|----------------------|--|---|
| Trypsin | R_2 not Pro $\text{R}_1 = \text{Lys, Arg}$ | From digestive systems of animals, many other sources |
| Chymotrypsin | R_2 not Pro $\text{R}_1 = \text{Tyr, Trp, Phe, Leu}$ | Same as trypsin |
| Thrombin | $\text{R}_1 = \text{Arg}$ | From blood; involved in coagulation |
| V-8 protease | $\text{R}_1 = \text{Asp, Glu}$ | From <i>Staphylococcus aureus</i> |
| Prolyl endopeptidase | $\text{R}_1 = \text{Pro}$ R_2 not Pro | Lamb kidney, other tissues |
| Subtilisin | Very little specificity | From various bacilli |
| Carboxypeptidase A | $\text{R}_2 = \text{C-terminal amino acid}$ | From digestive systems of animals |
| Thermolysin | R_1 not Pro $\text{R}_2 = \text{Leu, Val, Ile, Met}$ | From <i>Bacillus thermoproteolyticus</i> |
| Cyanogen bromide | $\text{R}_1 = \text{Met}$ | |

^aThe residues indicated are those next to which cleavage is most likely. In some cases, preference is determined by the residue on the N-terminal side of the cleaved bond (R_1) and sometimes by the residue to the C-terminal side (R_2). Generally, proteases do not cleave where proline is on the other side of the bond. Even prolyl endopeptidase will not cleave if $\text{R}_2 = \text{Pro}$.



Edman Degradation for N-terminal sequencing of peptides

- Repeated proteolysis
- For step-wise N-terminal identification: gives 1st aa, 2nd aa, 3rd aa etc.

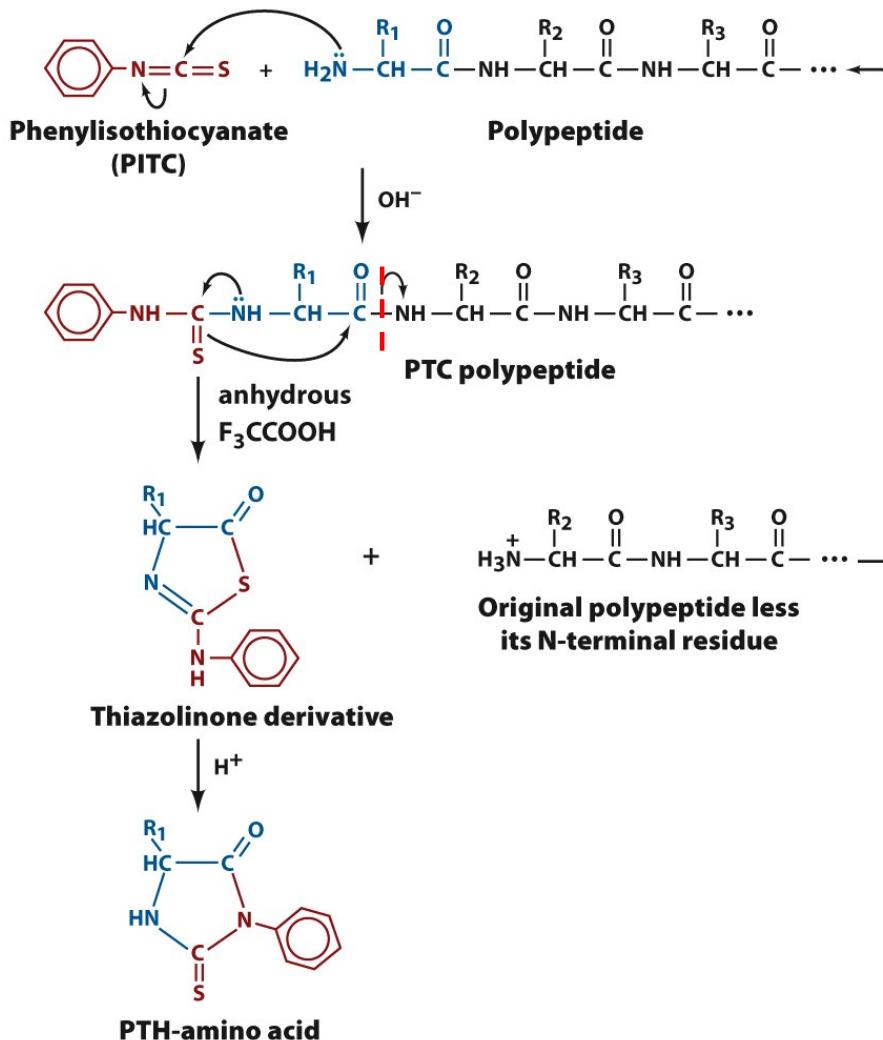
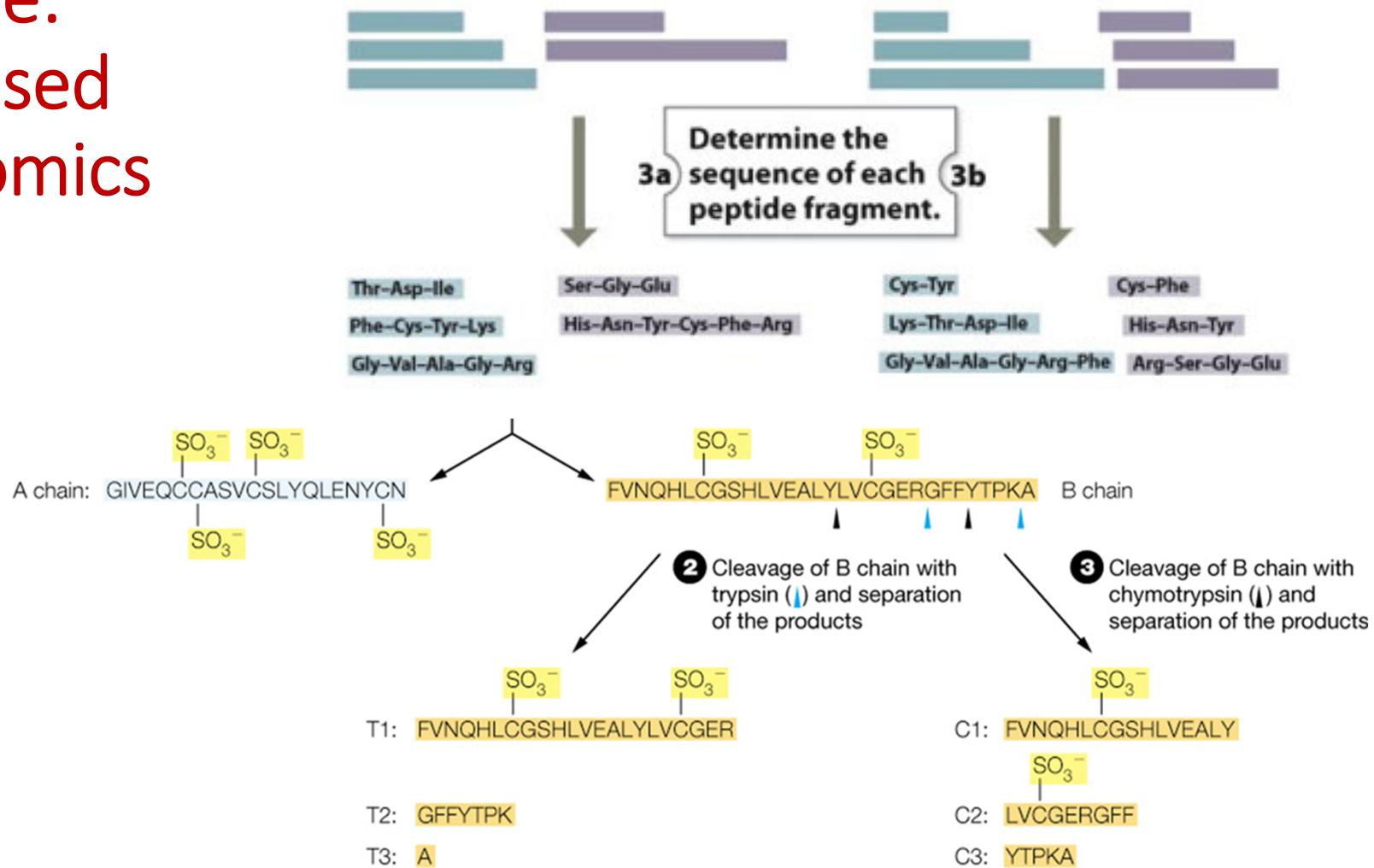


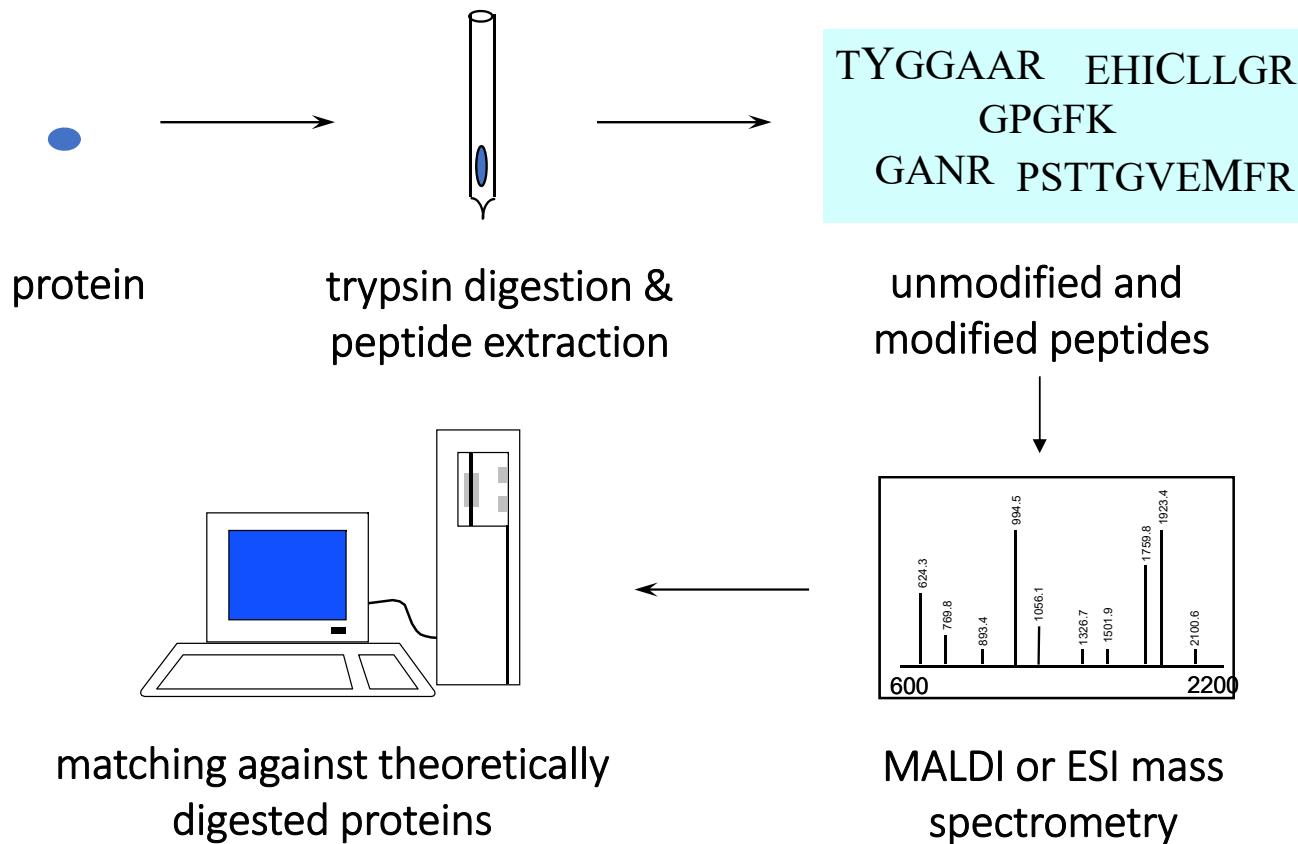
Figure 5-16
© 2013 John Wiley & Sons, Inc. All rights reserved.

Step 3: Determining the primary sequence of each peptide: MS-based Proteomics

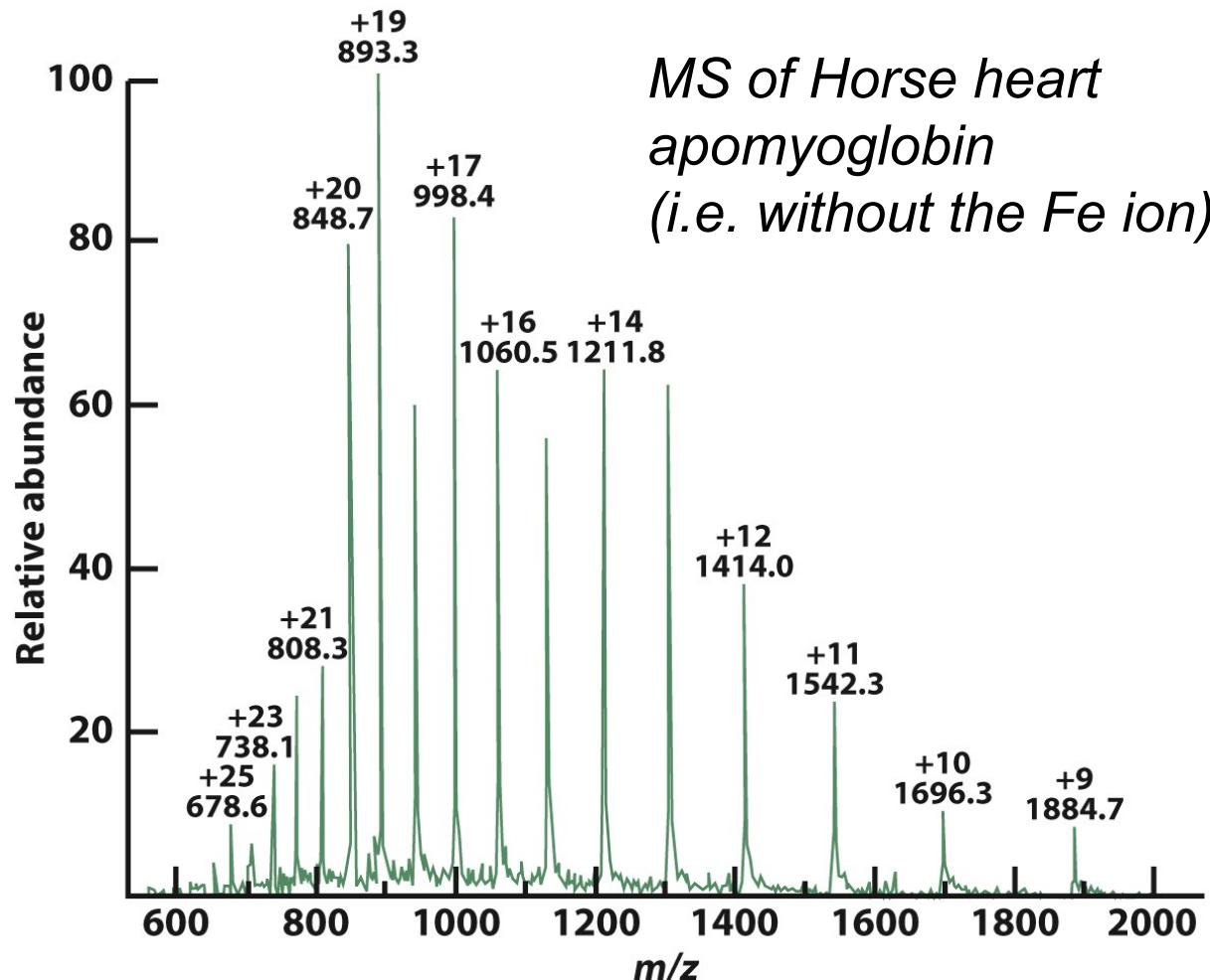


Peptide Masses from Mass Spectrometry

Mass spectrometers accurately measure peptide masses

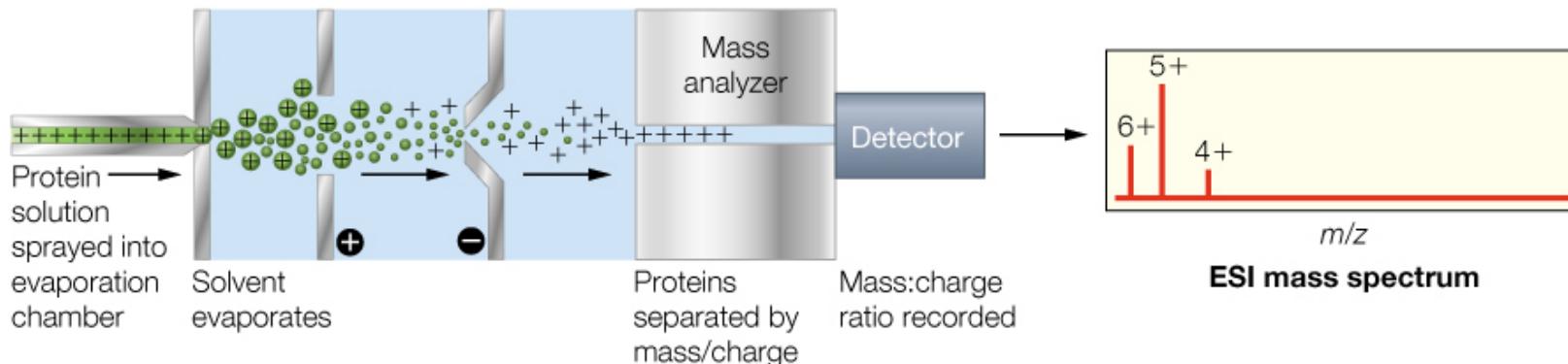


Mass Spectrum gives m/z of peptides



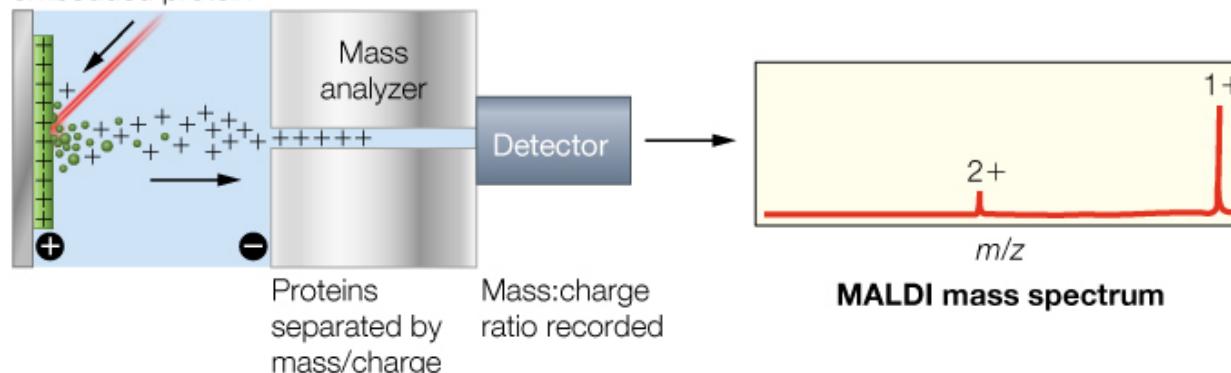
MS methods

Electrospray Ionization (ESI)



Matrix-Assisted Laser Desorption/Ionization (MALDI)

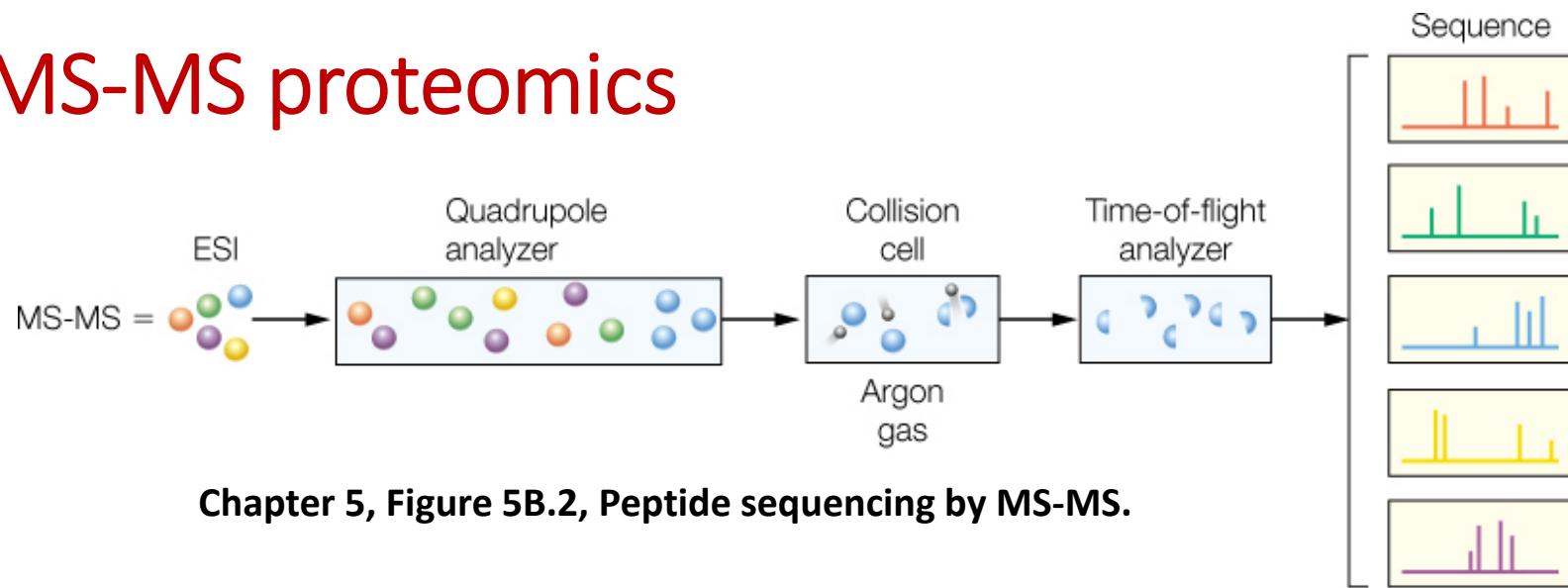
Laser pulse (orange) vaporizes matrix containing embedded protein



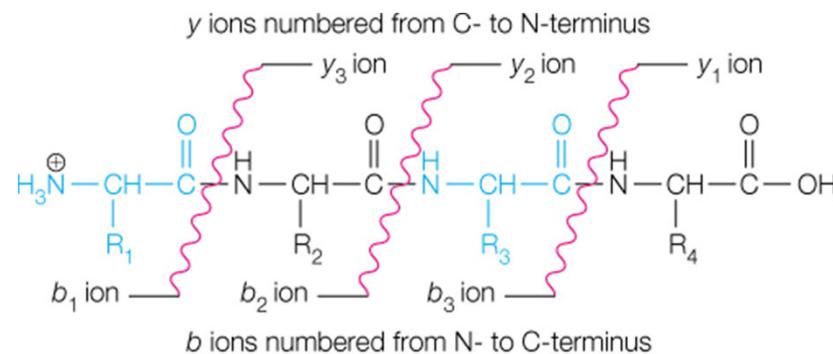
Chapter 5, Figure 5B.1, Electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI) mass spectrometry techniques.



MS-MS proteomics



Chapter 5, Figure 5B.2, Peptide sequencing by MS-MS.



Chapter 5, Figure 5B.3, Principal ions generated by low-energy collision-induced fragmentation.

Sequencing of the B Chain of Insulin Using Proteases and MS-MS: locating disulphide bonds in insulin

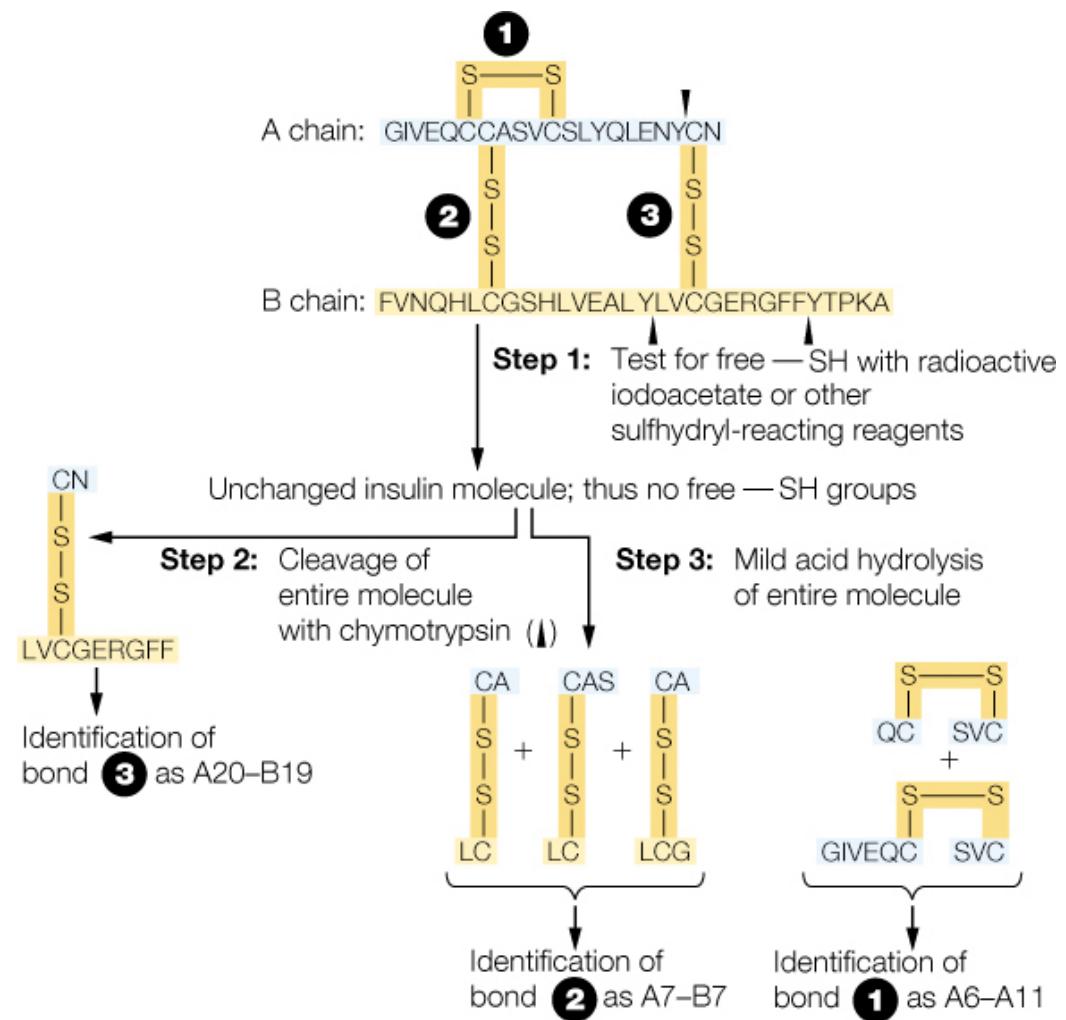
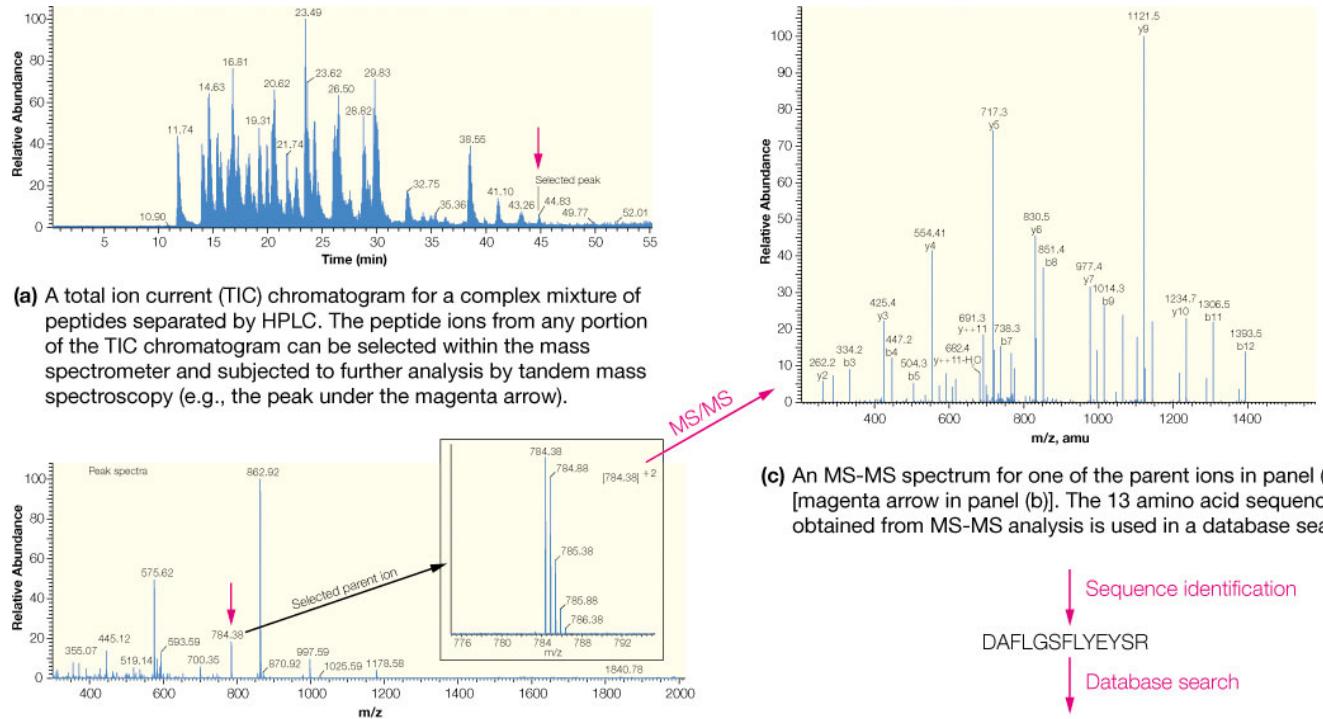
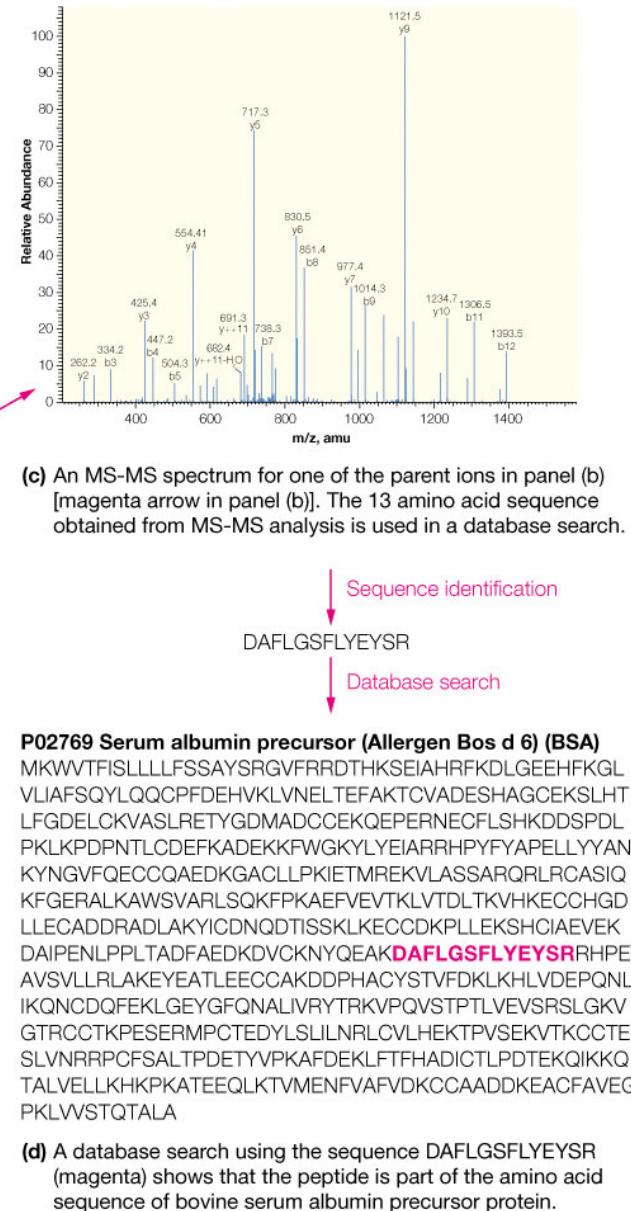


FIGURE 5B.5 Locating the disulfide bonds in insulin.

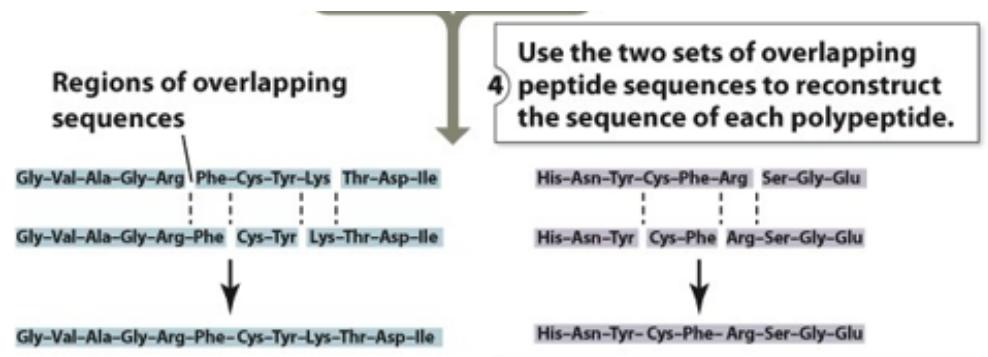
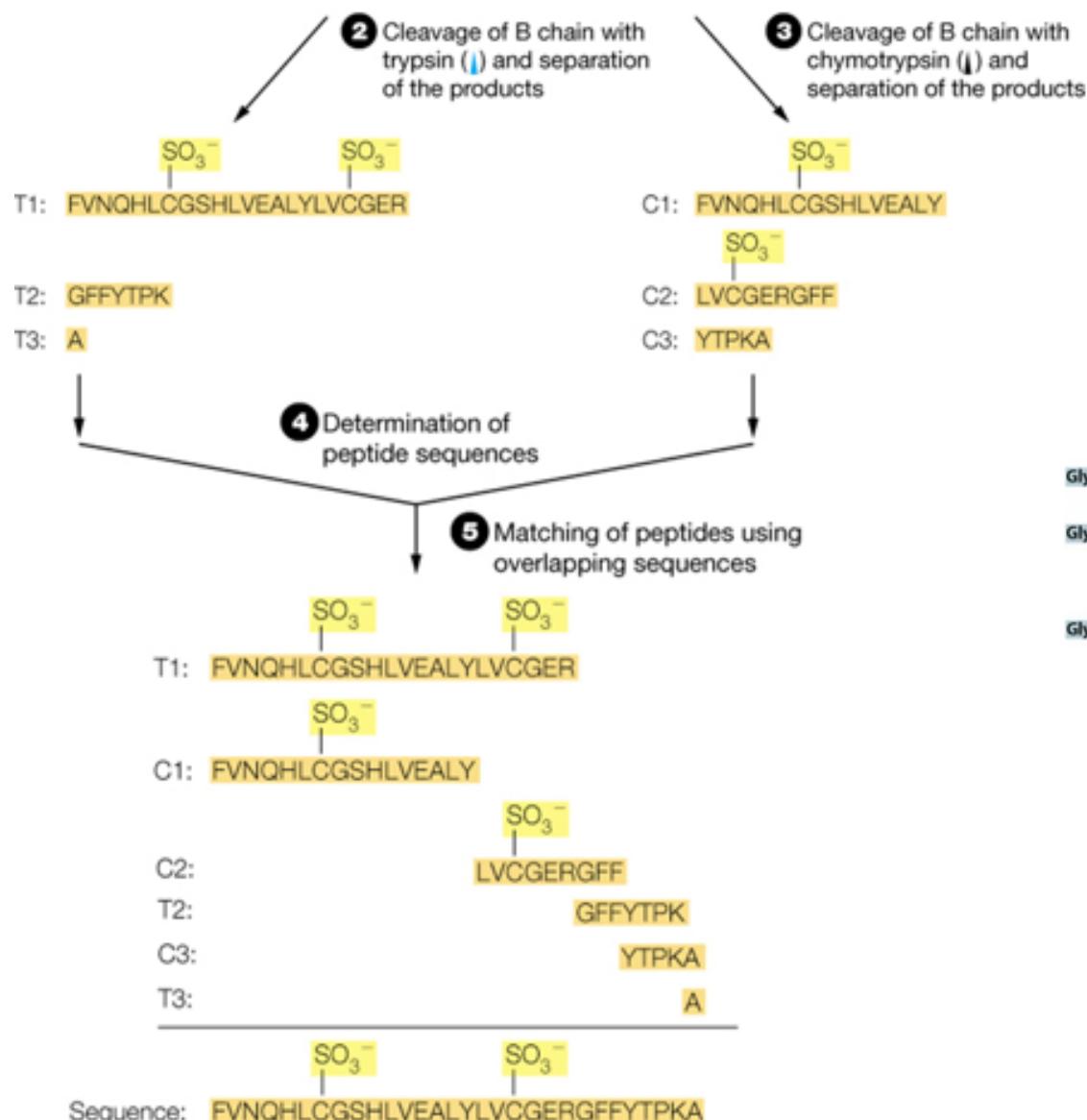
Finding possible proteins from proteomics



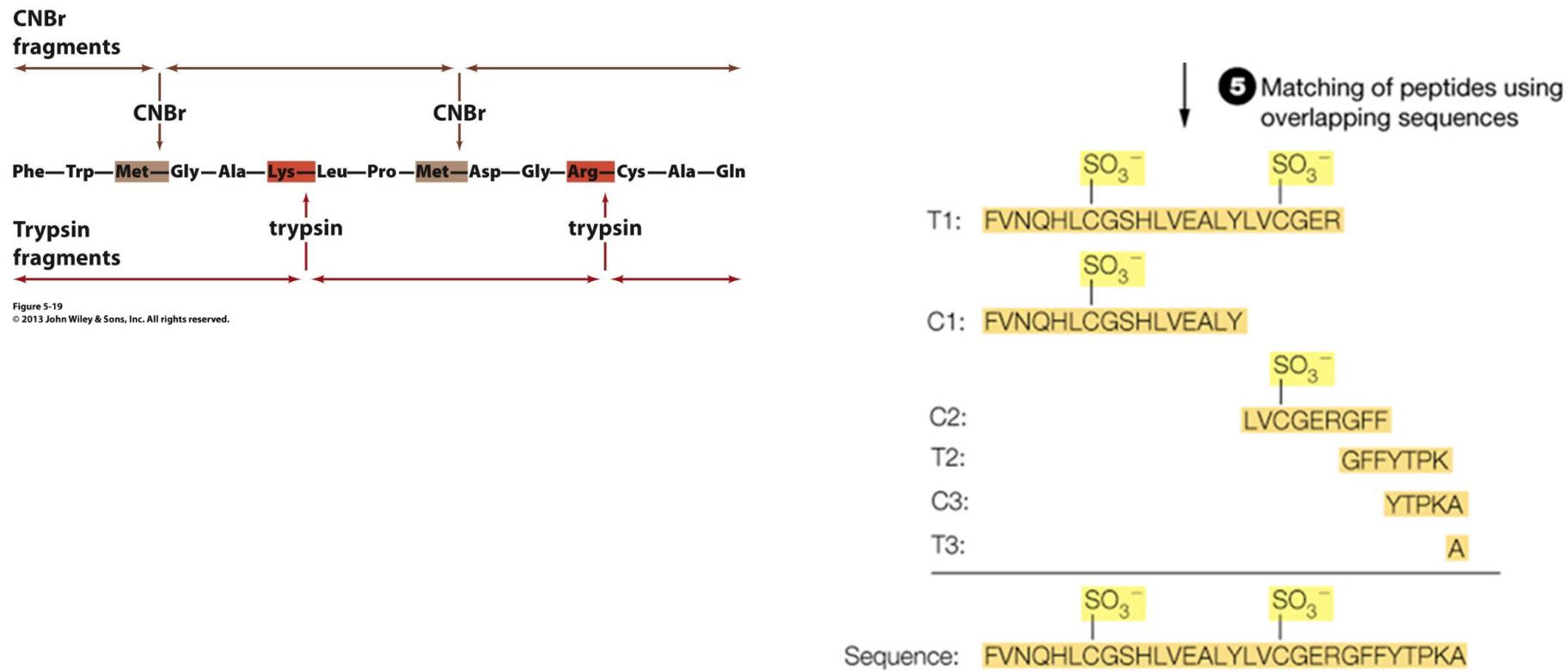
Chapter 5, Figure 5B.6, Identification of a protein of interest using proteomics methods.



Step 4: Figuring out how to fit the peptide sequences together



Overlapping Fragments Allow Determination of Amino Acid Sequence



Protein & DNA Sequence Data Banks for identifying peptides from proteomics

TABLE 5-5 Internet Addresses for the Major Protein and DNA Sequence Data Banks

Data Banks Containing Protein Sequences

ExPASy Proteomics Server: <http://expasy.org/>

Protein Information Resource (PIR): <http://pir.georgetown.edu/>

UniProt: <http://www.uniprot.org/>

Data Banks Containing Gene Sequences

GenBank: <http://www.ncbi.nlm.nih.gov/genbank/>

European Bioinformatics Institute (EBI): <http://www.ebi.ac.uk>

GenomeNet: [http://www.genome.jp/](http://www.genome.jp)

Table 5-5
© 2013 John Wiley & Sons, Inc. All rights reserved.



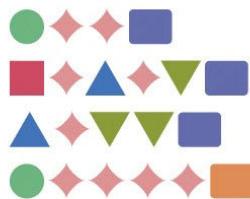
Protein sequences reveal shared parts called domains

(a) Fibronectin



(b) Blood clotting proteins

Factors VII, IX, X, and protein C



Factor XII

Tissue-type plasminogen activator

Protein S

Key

- ▲ Fibronectin domain 1
- Fibronectin domain 2
- Fibronectin domain 3
- γ-Carboxyglutamate domain
- ◆ Epidermal growth factor domain
- Serine protease domain
- ▼ Kringle domain
- Unique domain

1. Parts that are common to different proteins are called **domains**.
2. A single domain usually has the same function, even in different proteins.
3. In combination with other domains, proteins have different overall functionality.

Protein Sequencing summary

- To be sequenced, a protein must be separated into individual polypeptides that can be cleaved into sets of overlapping fragments.
- The amino acid sequence can be determined by Edman degradation, a procedure for removing N-terminal residues one at a time – very slow.
- Mass spectrometry can quickly identify amino acid sequences from the mass-to-charge ratio of gas-phase protein fragments.
- Protein sequence data are deposited in online databases.

Points to think about

- Summarize the steps involved in sequencing a protein.
- Why is it important to identify the N-terminal residue(s) of a protein?
- What are some advantages of sequencing peptides by mass spectrometry rather than by Edman degradation?
- Explain why long polypeptides must be broken into at least two different sets of peptide fragments for sequencing.
- What types of information can be retrieved from a protein sequence database?

Protein structure – definition of terms

- Zeroth level of protein structure: aa composition (historical)
- Primary (1°) structure
 - Amino acid sequence
- Secondary (2°) structure
 - Structural elements in proteins that are primarily formed through peptide backbone hydrogen bonds
- Tertiary (3°) structure
 - The three dimensional (3D) structure of a protein
- Quaternary (4°) Structure
 - Arrangement of subunits within a multisubunit protein

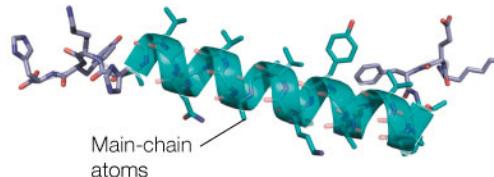
1° , 2° , 3° & 4° Structure: haemoglobin

...KEFTPPVQAAYQKVAGVANALAHKYH...

(a) Primary structure (amino acid sequence):

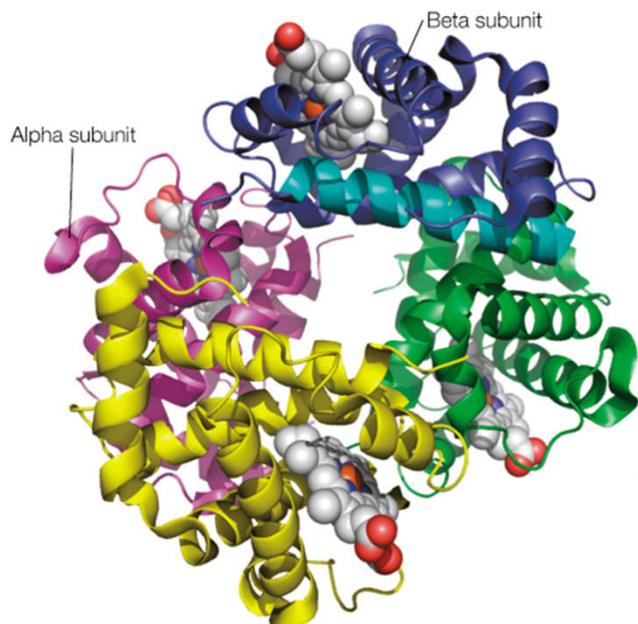
A portion of the amino acid sequence of human beta globin is shown. The sequence highlighted in cyan adopts a helical conformation, and is shown in the same orientation in parts (b-d).

Some parts of the primary sequence adopt a local regular repeating structure (" 2° structure")



(b) Secondary structure:

A stick representation of the amino acid sequence from part (a) is shown. Superimposed on the stick structure is a cartoon rendering of

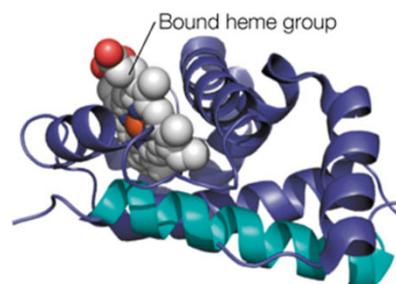


(d) Quaternary structure:

Four separate protein subunits, two alpha subunits (magenta and green) and two beta subunits (yellow and blue/cyan) associate to form the fully assembled hemoglobin protein. The four subunits are shown in cartoon rendering with hemes in space-filling display (PDB ID: 2hhb).

Quaternary structure (" 4° structure") arises when two or more proteins folded into tertiary structures interact to form well-defined multisubunit complexes.

Several 2° structure elements associate along their hydrophobic surfaces to give a stably folded structure (" 3° structure")



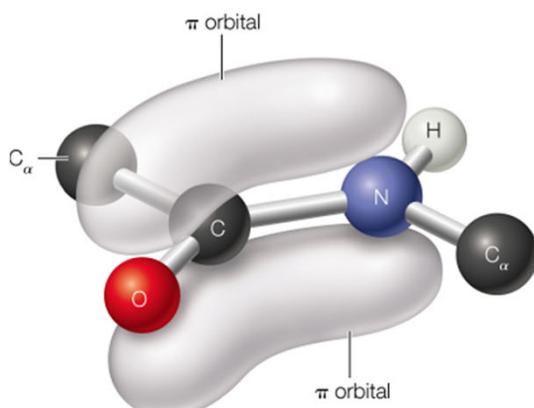
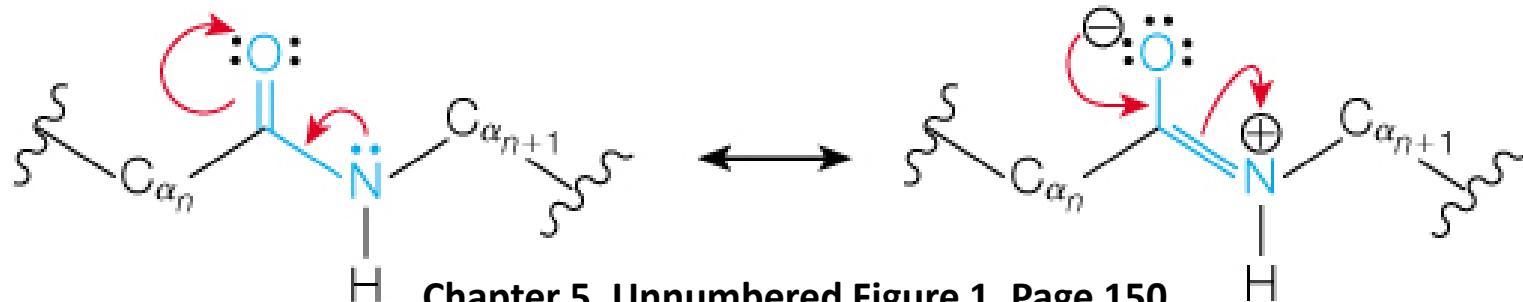
(c) Tertiary structure:

The entire beta globin chain is shown in its well-defined folded structure. As in myoglobin the helical regions interact to define the folded structure, which binds a heme (shown in space-filling display).

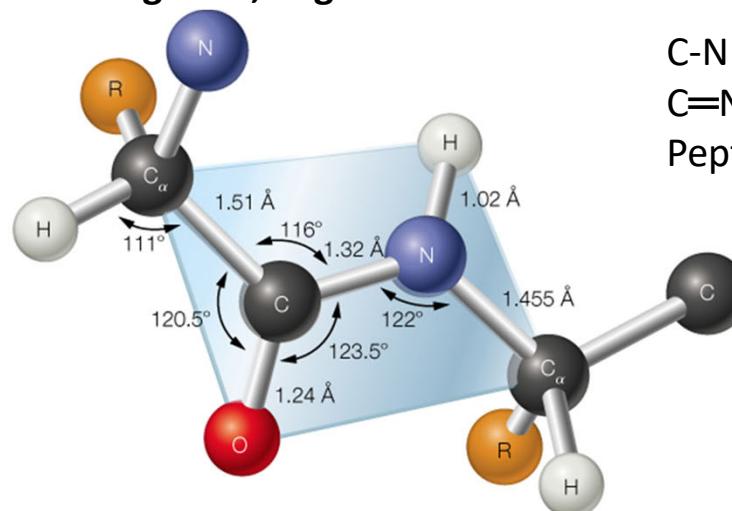
FIGURE 6.2 The four levels of structural organization in proteins.



Revisiting the peptide bond: neither single nor double



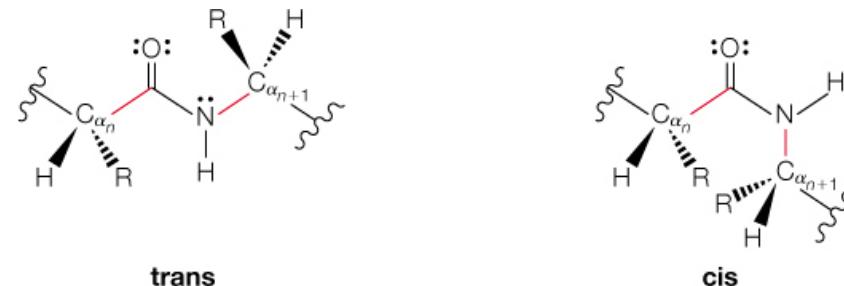
(a) Delocalization of the π -electron orbitals over the three atoms O—C—N accounts for the partial double-bond character of the C—N bond.



(b) The presently accepted values for bond angles and bond lengths (in Angstrom, Å or m^{-10}) are given here. The six atoms shown in (a) and those in the blue rectangle in (b) are nearly coplanar.

C—N single bond: 1.49 Å
 C=N double bond: 1.27 Å
 Peptide bond: 1.32 Å

Secondary structure: Peptide bonds impose planarity on the 6 atoms involved



Chapter 5, Unnumbered Figure 2, Page 150

Extended Conformation of Polypeptide

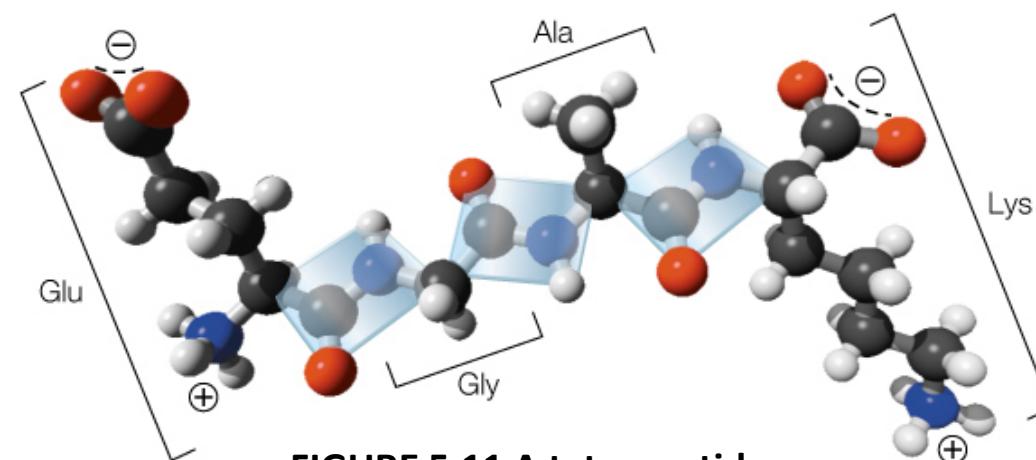


FIGURE 5.11 A tetrapeptide.



Rotations around the N-C_α and the C_α-C bonds of the protein backbone

- Free rotation is only allowed about the C_α carbons.
 - However, this rotation is restricted by steric interactions.
- Such rotation is described by two torsional angles termed ϕ (between N and C_α) and ψ (between C_α and the carbonyl C).
- Torsions refer to two different residues.
- Each residue maintains its peptide bond planarity of the atoms involved in the peptide bond.
- Steric interference of adjacent R groups leads to only some allowed Φ (phi) and Ψ (psi) values.

φ (phi), ψ (psi)

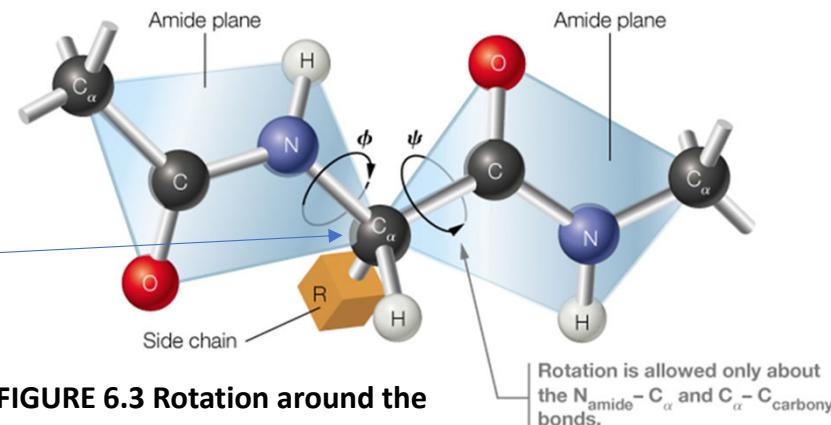
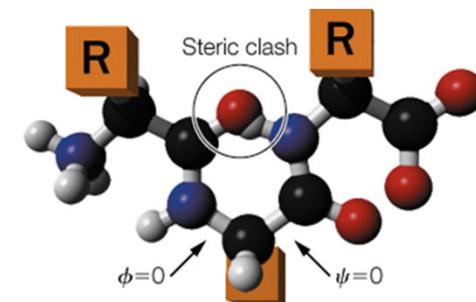


FIGURE 6.3 Rotation around the bonds in a polypeptide backbone.



(a) A sterically nonallowed conformation. The conformation $\phi = 0^\circ$, $\psi = 0^\circ$ is not allowed in any polypeptide chain because of the steric crowding between main-chain atoms. A tripeptide is shown, where the central amino acid has $\phi = 0^\circ$ and $\psi = 0^\circ$. Notice that the carbonyl oxygen of residue #1 (on the left) would clash with the amide hydrogen of residue #3 (on the right).

Figure 6.8, Steric interactions determine peptide conformation.

Naturally occurring secondary structures in proteins

Figure 6.4

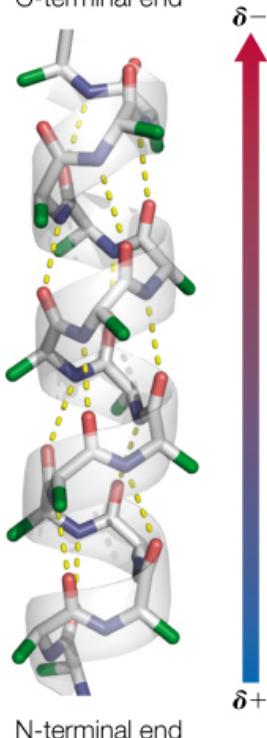
The right-handed α helix,

β sheet,

and

3_{10} helix.

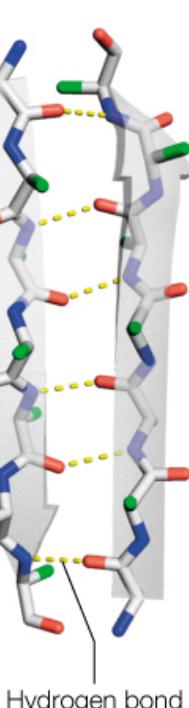
C-terminal end



N-terminal end

(a)

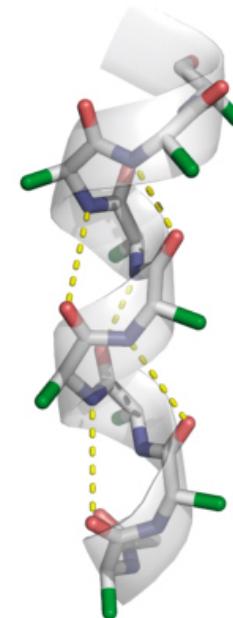
In the α helix, the hydrogen bonds are within a contiguous stretch of amino acids and are almost parallel to the helix axis. This orientation of the amide bonds in the helix gives rise to a helical macrodipole moment shown by the arrow (see Figure 2.5). The N-terminal end of the helix has partial (+) charge character, and the C-terminal end has partial (−) charge character.



Hydrogen bond

(b)

In the β sheet, the hydrogen bonds are between adjacent strands (only two strands are shown here), which are not necessarily contiguous in the primary sequence. In this structure, the hydrogen bonds are nearly perpendicular to the chains. Note that in the cartoon rendering a strand is shown by a flat arrow, where the head of the arrow points to the C-terminus of the strand.



(c)

The 3_{10} helix is found in proteins but is less common than the α helix. Note that, compared to the α helix, the 3_{10} helix forms a tighter spiral.



Summary

- The peptide bond imposes planarity on the 6 backbone atoms of each peptide
 - Free rotation is only allowed about the C_α carbons.
- Steric interactions from side chains limit the types of secondary structures that can form in 3D space.
 - Secondary structures are mainly stabilized by backbone hydrogen bonds.
 - There are two main types of structures that are able to form these bonds
 - ❖ the **α -helix** and
 - ❖ the **β -sheet**, made up of β -strands.

Spots test on textbook website

- Access textbook via iLearn
- Look for L5 Spot Test
- 2 questions; 5 mins; one attempt

Proteins 3: 2° & 3° Structure & Folding and Stability

Shoba Ranganathan

Applies Biosciences

T: 02 9850 6262; E: shoba.ranganathan@mq.edu.au

Outline

- Secondary Structure: Regular Ways to Fold the Polypeptide Chain
- Fibrous Proteins: Structural Materials of Cells and Tissues
- Globular Proteins: Tertiary Structure and Functional Diversity
- Factors Determining Secondary and Tertiary Structure
- Dynamics of Globular Protein Structure
- Prediction of Protein Secondary and Tertiary Structure
- Quaternary Structure of Proteins

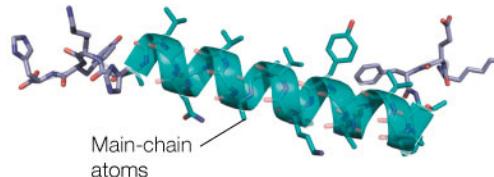
1° , 2° , 3° & 4° Structure: haemoglobin

...KEFTPPVQAAYQKVAGVANALAHKYH...

(a) Primary structure (amino acid sequence):

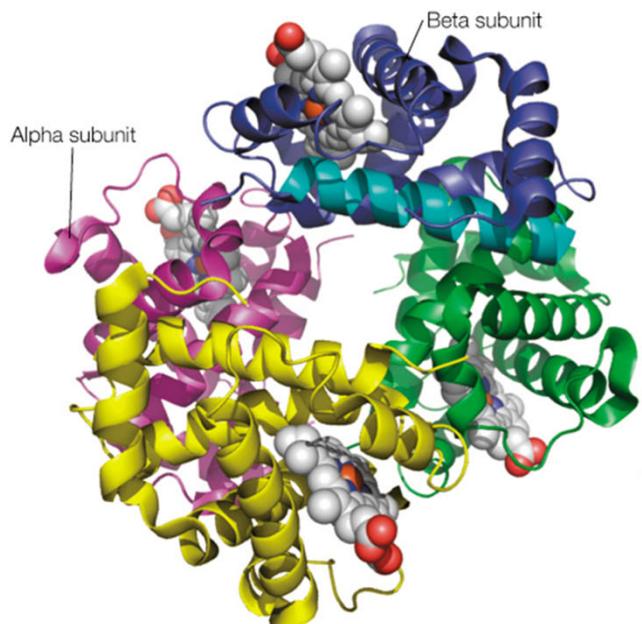
A portion of the amino acid sequence of human beta globin is shown. The sequence highlighted in cyan adopts a helical conformation, and is shown in the same orientation in parts (b-d).

Some parts of the primary sequence adopt a local regular repeating structure (" 2° structure")



(b) Secondary structure:

A stick representation of the amino acid sequence from part (a) is shown. Superimposed on the stick structure is a cartoon rendering of

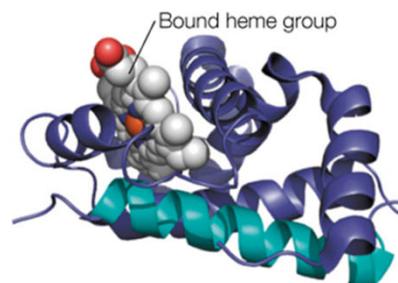


(d) Quaternary structure:

Four separate protein subunits, two alpha subunits (magenta and green) and two beta subunits (yellow and blue/cyan) associate to form the fully assembled hemoglobin protein. The four subunits are shown in cartoon rendering with hemes in space-filling display (PDB ID: 2hhb).

Quaternary structure (" 4° structure") arises when two or more proteins folded into tertiary structures interact to form well-defined multisubunit complexes.

Several 2° structure elements associate along their hydrophobic surfaces to give a stably folded structure (" 3° structure")



(c) Tertiary structure:

The entire beta globin chain is shown in its well-defined folded structure. As in myoglobin the helical regions interact to define the folded structure, which binds a heme (shown in space-filling display).

FIGURE 6.2 The four levels of structural organization in proteins.



Naturally occurring secondary structures in proteins

Figure 6.4

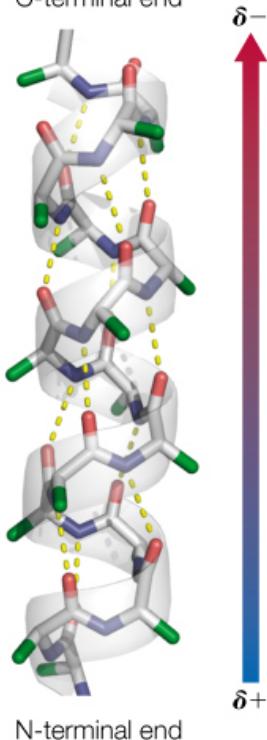
The right-handed α helix,

β sheet,

and

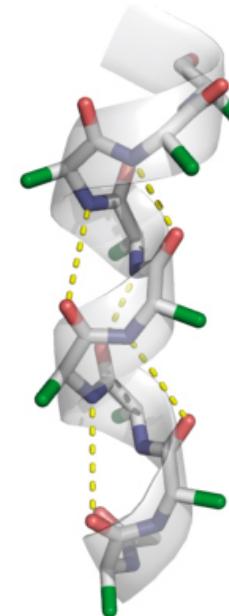
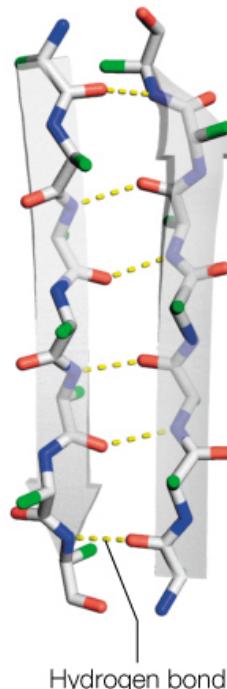
3_{10} helix.

C-terminal end



N-terminal end

- (a) In the α helix, the hydrogen bonds are within a contiguous stretch of amino acids and are almost parallel to the helix axis. This orientation of the amide bonds in the helix gives rise to a helical macrodipole moment shown by the arrow (see Figure 2.5). The N-terminal end of the helix has partial (+) charge character, and the C-terminal end has partial (−) charge character.
- (b) In the β sheet, the hydrogen bonds are between adjacent strands (only two strands are shown here), which are not necessarily contiguous in the primary sequence. In this structure, the hydrogen bonds are nearly perpendicular to the chains. Note that in the cartoon rendering a strand is shown by a flat arrow, where the head of the arrow points to the C-terminus of the strand.



- (c) The 3_{10} helix is found in proteins but is less common than the α helix. Note that, compared to the α helix, the 3_{10} helix forms a tighter spiral.



Common Secondary Structure Elements

Side chain positions in an α -helix

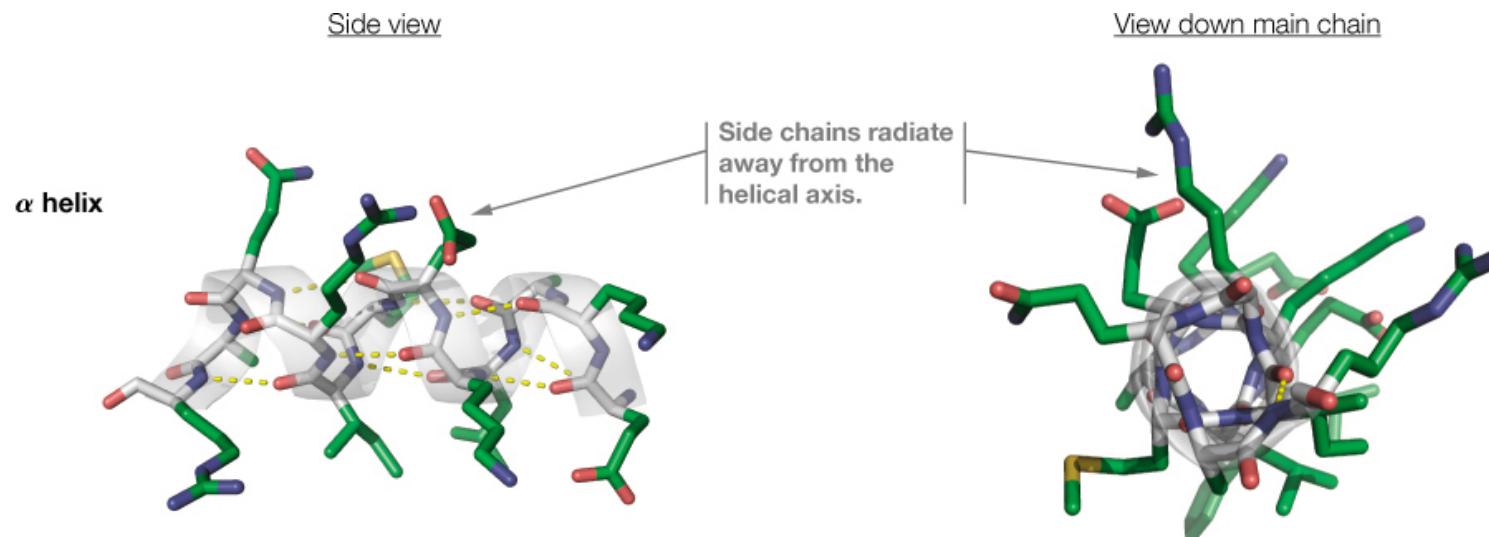


FIGURE 6.7 The positions of side chains in the α helix and β sheet.

In an α -helix, the side chains radiate away from the helical axis. Its center consists backbone atoms, closely packed. The hydrogen bonds that stabilize the helix are shown as yellow dashes.

Common Secondary Structure Elements

Side chain positions in a β -sheet

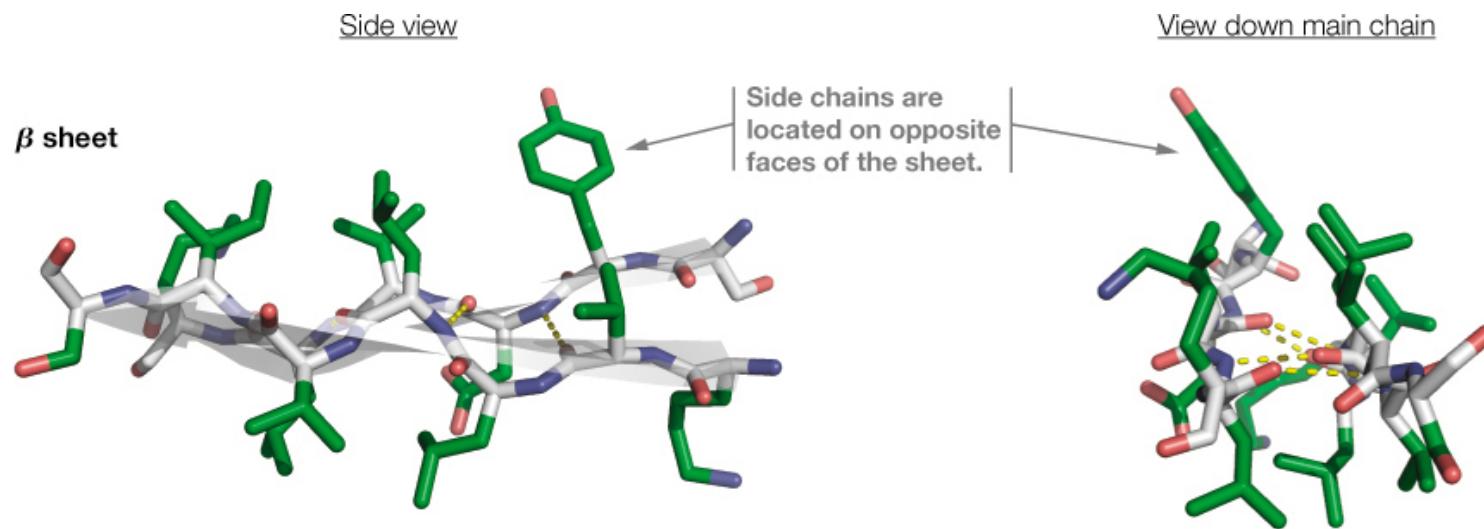


FIGURE 6.7 The positions of side chains in the α helix and β sheet.

In a β -sheet, neighboring side chains are located on opposite faces of the sheet, which is stabilized by main-chain hydrogen bonds between adjacent β -strands.

Rotations around the N-C_α and the C_α-C bonds of the protein backbone

- Free rotation is only allowed about the C_α carbons.
 - However, this rotation is restricted by steric interactions.
- Such rotation is described by two torsional angles termed ϕ (between N and C_α) and ψ (between C_α and the carbonyl C).
- Torsions refer to two different residues.
- Each residue maintains its peptide bond planarity of the atoms involved in the peptide bond.
- Steric interference of adjacent R groups leads to only some allowed Φ (phi) and Ψ (psi) values.

φ (phi), ψ (psi)

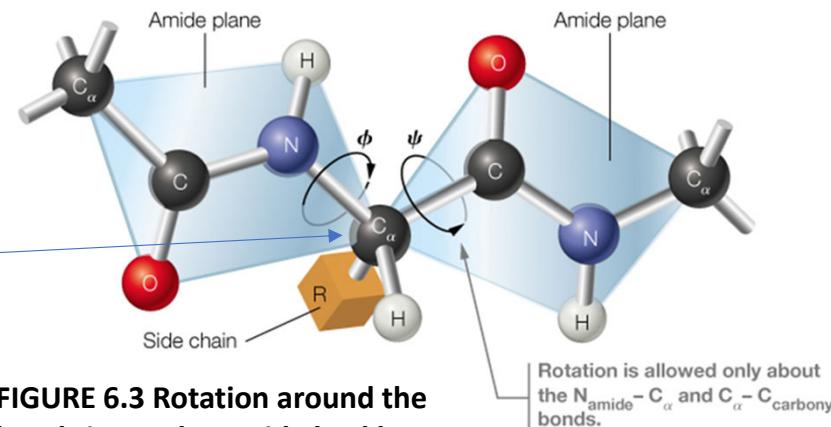
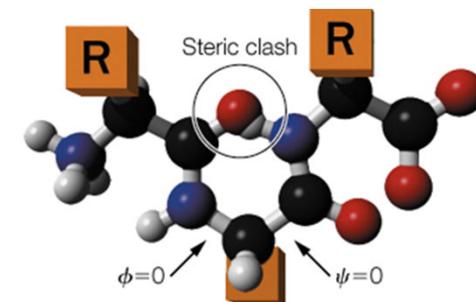


FIGURE 6.3 Rotation around the bonds in a polypeptide backbone.

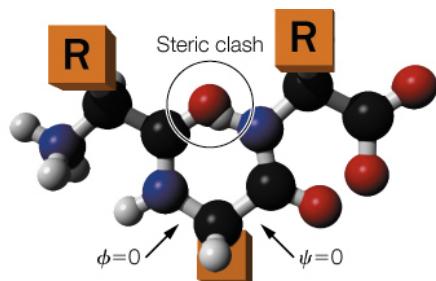


(a) A sterically nonallowed conformation. The conformation $\phi = 0^\circ$, $\psi = 0^\circ$ is not allowed in any polypeptide chain because of the steric crowding between main-chain atoms. A tripeptide is shown, where the central amino acid has $\phi = 0^\circ$ and $\psi = 0^\circ$. Notice that the carbonyl oxygen of residue #1 (on the left) would clash with the amide hydrogen of residue #3 (on the right).

Figure 6.8, Steric interactions determine peptide conformation.

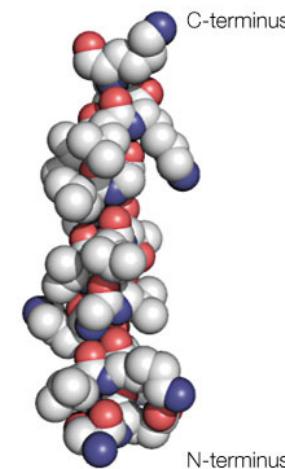
Common Secondary Structure Elements

Steric Interactions determine peptide conformation



(a) A sterically nonallowed conformation. The conformation $\phi = 0^\circ$, $\psi = 0^\circ$ is not allowed in any polypeptide chain because of the steric crowding between main-chain atoms. A tripeptide is shown, where the central amino acid has $\phi = 0^\circ$ and $\psi = 0^\circ$. Notice that the carbonyl oxygen of residue #1 (on the left) would clash with the amide hydrogen of residue #3 (on the right).

Certain ϕ and ψ angles result in steric clashes, where atoms are closer than their van der Waals radii. These angles and related conformations are not allowed.



(b) The atoms in a helix are closely packed but do not clash sterically. Here, a segment of an α helix in sperm whale myoglobin is shown as a space-filling model (this is the longer green helix in Figure 6.1; PDB ID: 1mbn).

The backbone of an α helix results in closely packed atoms that do not sterically clash.

FIGURE 6.8 Steric interactions determine peptide conformation.

Ramachandran Plots

Graphing sterically allowed ϕ and ψ angles

- poly-L-alanine

- white areas are from theoretical predictions
- Glycines can go anywhere

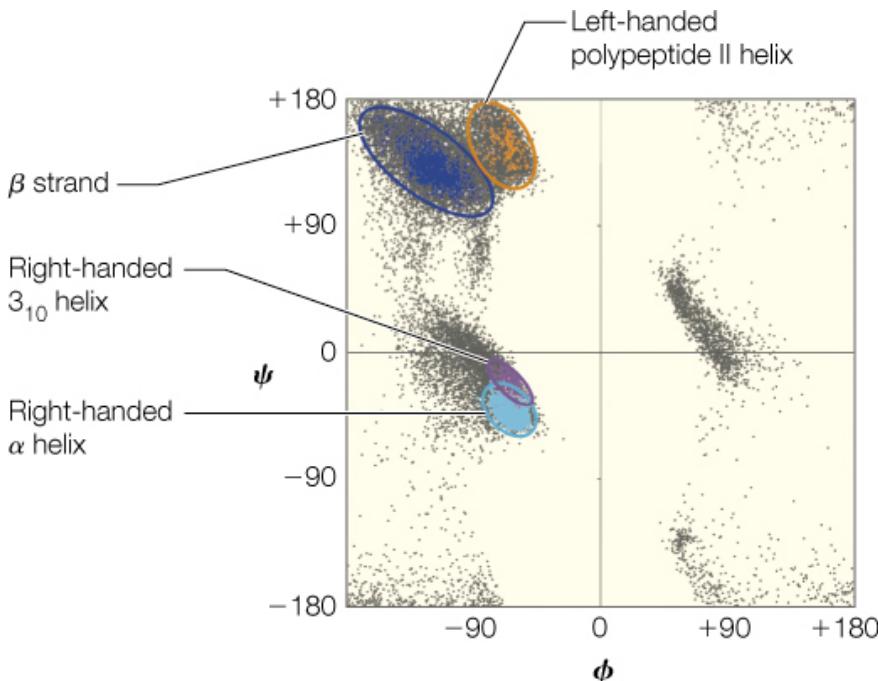


FIGURE 6.11 Observed values of ϕ and ψ from protein structural data.

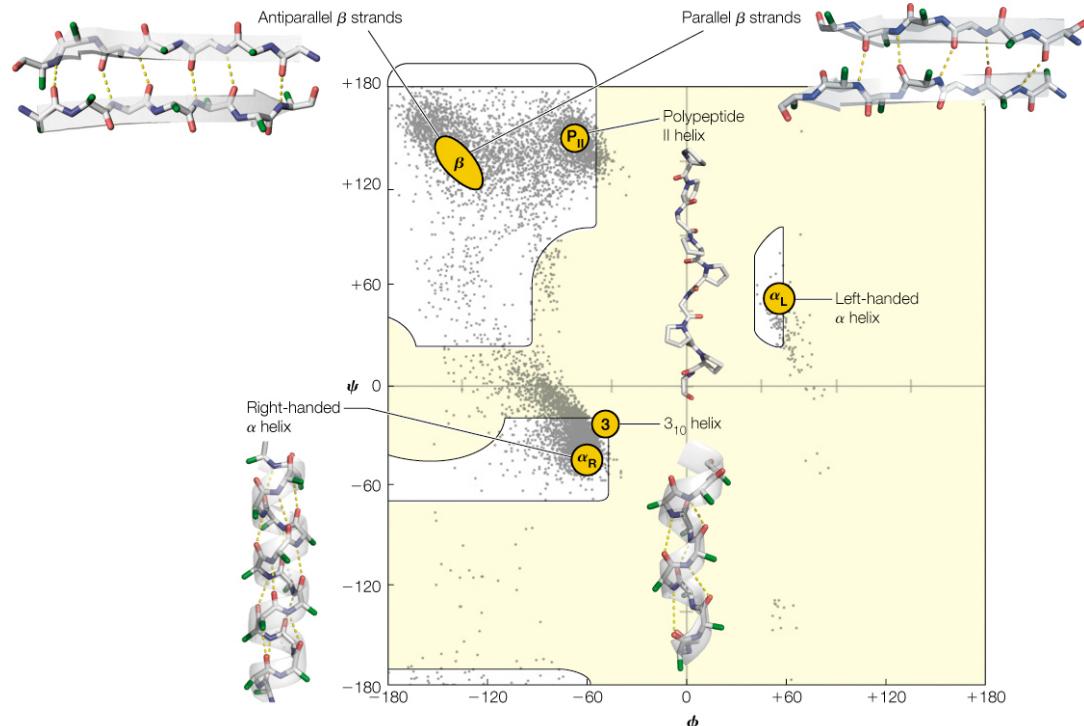


TABLE 6.2 Ranges of allowed ϕ and ψ angles for some polypeptide secondary structures

| Structure Type | ϕ | ψ |
|----------------------|------------------------------|------------------------------|
| β strand | -150° to -100° | $+120^\circ$ to $+160^\circ$ |
| α helix | -70° to -60° | -50° to -40° |
| 3_{10} helix | -70° to -60° | -30° to -10° |
| Polypeptide II helix | -80° to -60° | $+130^\circ$ to $+160^\circ$ |

Data from *Protein Science* 18:1321–1325 (2009), S. A. Hollingsworth, D. S. Berkholz, and P. A. Karplus, On the occurrence of linear groups in proteins.

FIGURE 6.9 A Ramachandran plot for poly-L-alanine.

Fibrous Proteins as Structural Materials

- Fibrous proteins are elongated molecules with well-defined secondary structures
- Examples include:
 - keratin—hair, fingernails, feathers, scales, or intermediate filaments (intracellular)
 - fibroin—silk cocoons
 - collagen—abundant connective tissue protein; matrix material in bone on which mineral components precipitate

- High abundance of certain amino acids in fibrous proteins

TABLE 6.3 Amino acid compositions of some fibrous proteins

| Amino Acid | α -Keratin (wool) | Fibroin (silk) | Collagen (bovine tendon) | All Proteins ^c |
|------------|--------------------------|----------------|--------------------------|---------------------------|
| Gly | 8.1 | 44.6 | 32.7 | 7.9 |
| Ala | 5.0 | 29.4 | 12.0 | 8.7 |
| Ser | 10.2 | 12.2 | 3.4 | 5.8 |
| Glu + Gln | 12.1 | 1.0 | 7.7 | 6.6 (3.7) |
| Cys | 11.2 | 0 | 0 | 1.3 |
| Pro | 7.5 | 0.3 | 22.1 ^a | 4.7 |
| Arg | 7.2 | 0.5 | 5.0 | 5.0 |
| Leu | 6.9 | 0.5 | 2.1 | 8.9 |
| Thr | 6.5 | 0.9 | 1.6 | 5.6 |
| Asp + Asn | 6.0 | 1.3 | 4.5 | 5.9 (4.2) |
| Val | 5.1 | 2.2 | 1.8 | 7.2 |
| Tyr | 4.2 | 5.2 | 0.4 | 3.5 |
| Ile | 2.8 | 0.7 | 0.9 | 5.5 |
| Phe | 2.5 | 0.5 | 1.2 | 4.0 |
| Lys | 2.3 | 0.3 | 3.7 ^b | 5.5 |
| Trp | 1.2 | 0.2 | 0 | 1.5 |
| His | 0.7 | 0.2 | 0.3 | 2.4 |
| Met | 0.5 | 0 | 0.7 | 2.0 |

Note: The three most abundant amino acids in each protein are indicated in magenta. Values are given in mole percent.

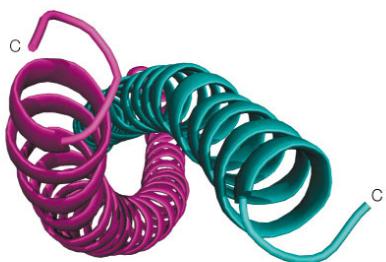
^aAbout 39% of this is hydroxyproline.

^bAbout 14% of this is hydroxylysine.

^cData from *Journal of Chemical Information and Modeling* (2010) 50:690–700, J. M. Otaki, M. Tsutsumi, T. Gotoh, and H. Yamamoto, Secondary structure characterization based on amino acid composition and availability in proteins.

Examples of fibrous proteins

In α -keratin, large hydrophobic residues repeat every four positions. The α -helix has 3.6 residues per turn, giving each helix a hydrophobic side, which defines the interface between two long helices in the coiled-coil structure typical of keratin.



(b) View looking down the axis of the coiled-coil from the C-terminal end of the two monomers (side chains removed for clarity).



(a) Side view of two monomers interacting via a parallel coiled-coil. N- and C-termini are indicated.

FIGURE 6.12 The coiled-coil structure of α -keratin intermediate filaments.

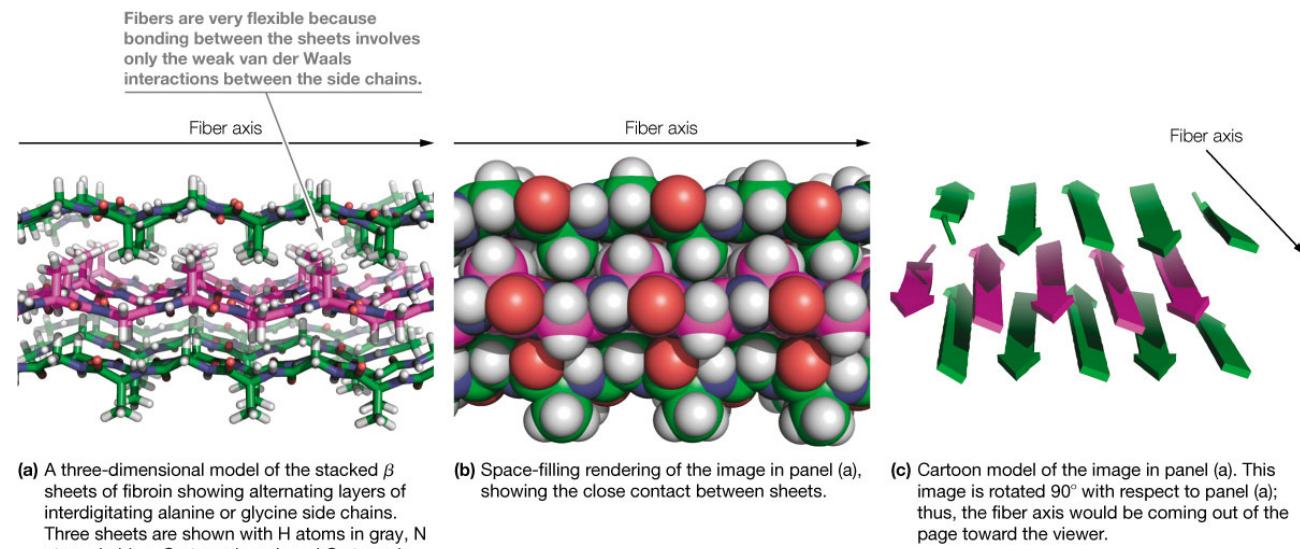
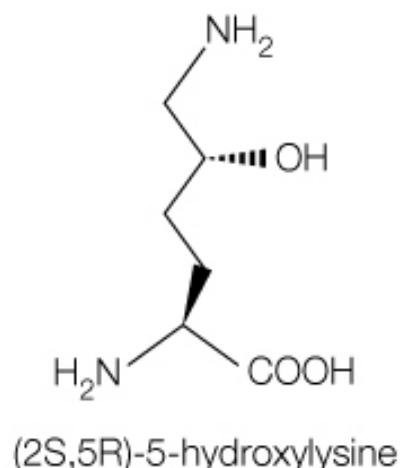
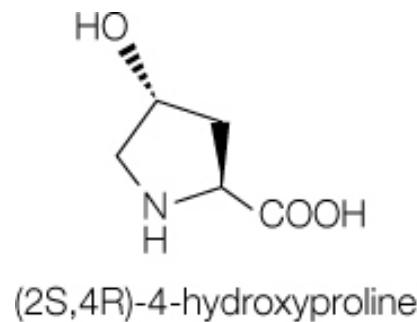


FIGURE 6.13 Theoretical model for the structure of silk fibroin.

The extensive close-packed β sheet of fibroin is interrupted by compact folded regions which provide some elasticity.

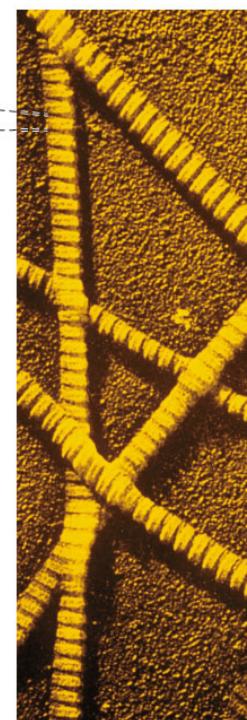
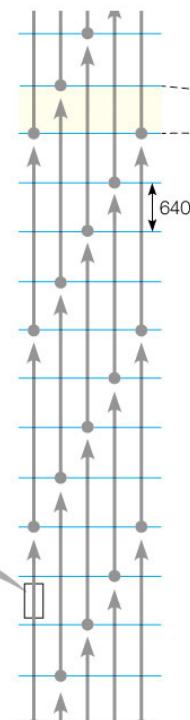
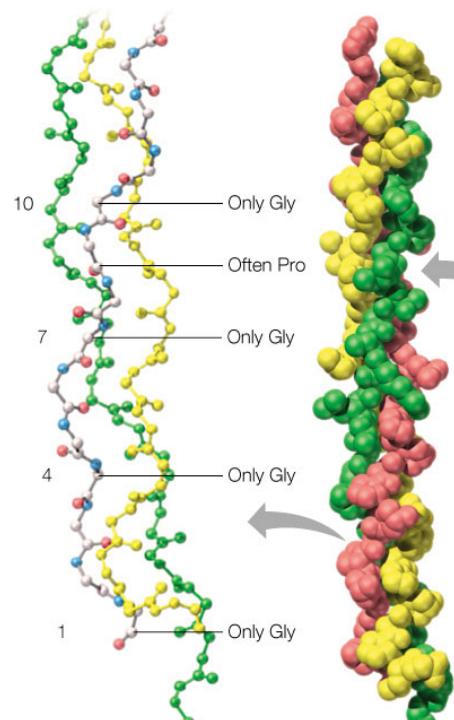
Collagen



- It is an abundant connective tissue protein; matrix material in bone, on which mineral components precipitate; triple-strand left-handed helix
- It contains hydroxyproline (Hyp) and hydroxylysine
- G-X-Y tripeptide motif, where X is Pro and Y is Pro or Hyp, lends itself to triple-strand structure
- Polypeptide chains are crosslinked and glycosylated
- Vitamin C (ascorbic acid) is a cofactor required for proline hydroxylation; vitamin C deficiency (scurvy) leads to collagen degeneration



Collagen

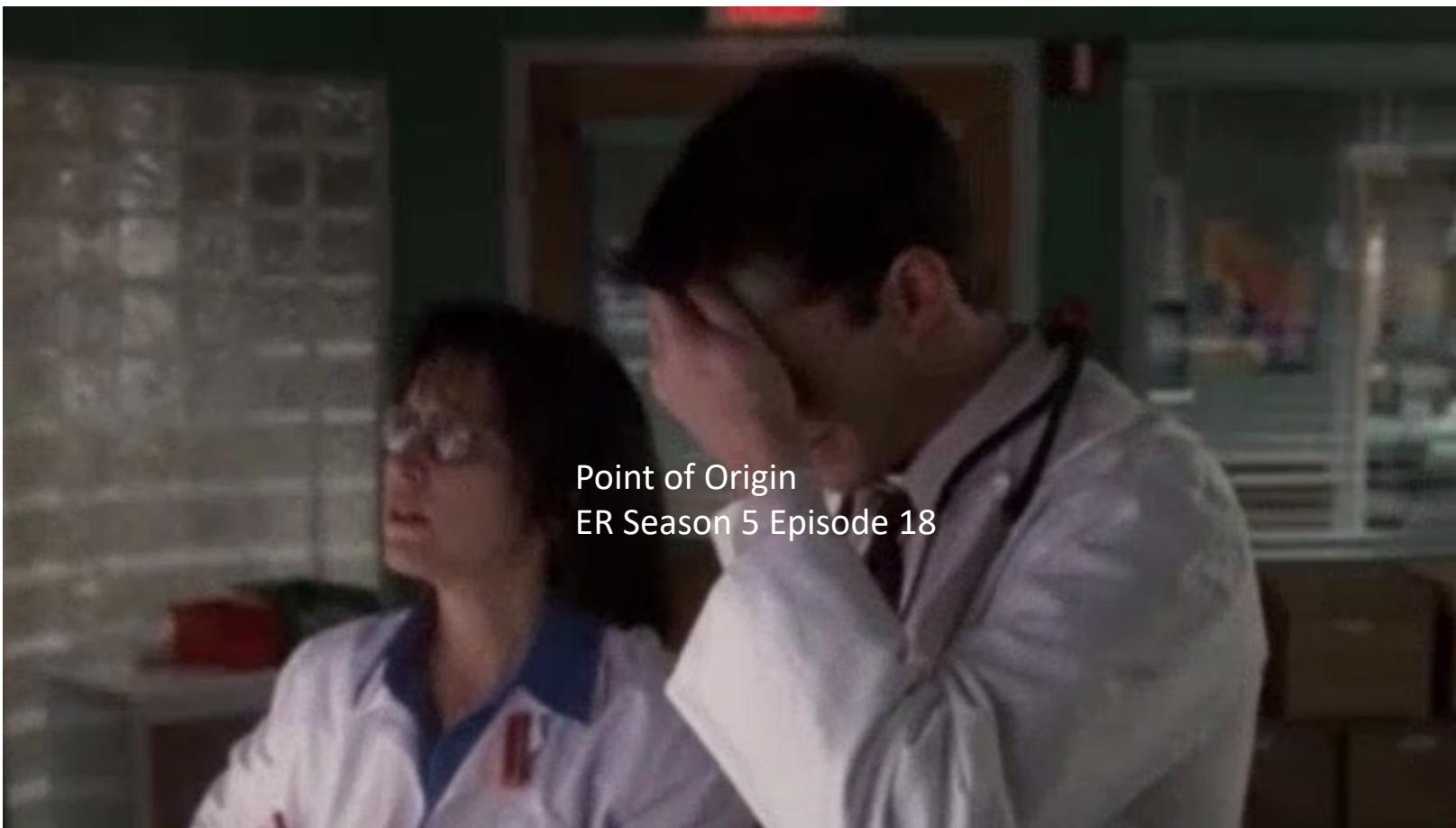


From an
electron
microscope

Collagen diseases

| | | |
|------------------|--------|---|
| Collagen disease | COL1: | Osteogenesis imperfecta · Ehlers–Danlos syndrome, types 1, 2, 7 |
| | COL2: | Hypochondrogenesis · Achondrogenesis type 2 · Stickler syndrome · Marshall syndrome · Spondyloepiphyseal dysplasia congenita · Spondyloepimetaphyseal dysplasia, Strudwick type · Kniest dysplasia (see also C2/11) |
| | COL3: | Ehlers–Danlos syndrome, types 3 & 4 (Sack–Barabas syndrome) |
| | COL4: | Alport syndrome |
| | COL5: | Ehlers–Danlos syndrome, types 1 & 2 |
| | COL6: | Bethlem myopathy · Ullrich congenital muscular dystrophy |
| | COL7: | Epidermolysis bullosa dystrophica · Recessive dystrophic epidermolysis bullosa · Bart syndrome · Transient bullous dermolysis of the newborn |
| | COL8: | Fuchs' dystrophy 1 |
| | COL9: | Multiple epiphyseal dysplasia 2, 3, 6 |
| | COL10: | Schmid metaphyseal chondrodysplasia |
| | COL11: | Weissenbacher–Zweymüller syndrome · Otospondylomegaepiphyseal dysplasia (see also C2/11) |
| | COL17: | Bullous pemphigoid |
| | COL18: | Knobloch syndrome |





Point of Origin
ER Season 5 Episode 18

Osteogenesis imperfecta:

<https://www.metacritic.com/tv/er/season-5/episode-18-point-of-origin>

Hydrogen bonds in an α -helix: from residue i to residue $i+4$: 1-5, 2-6, etc.

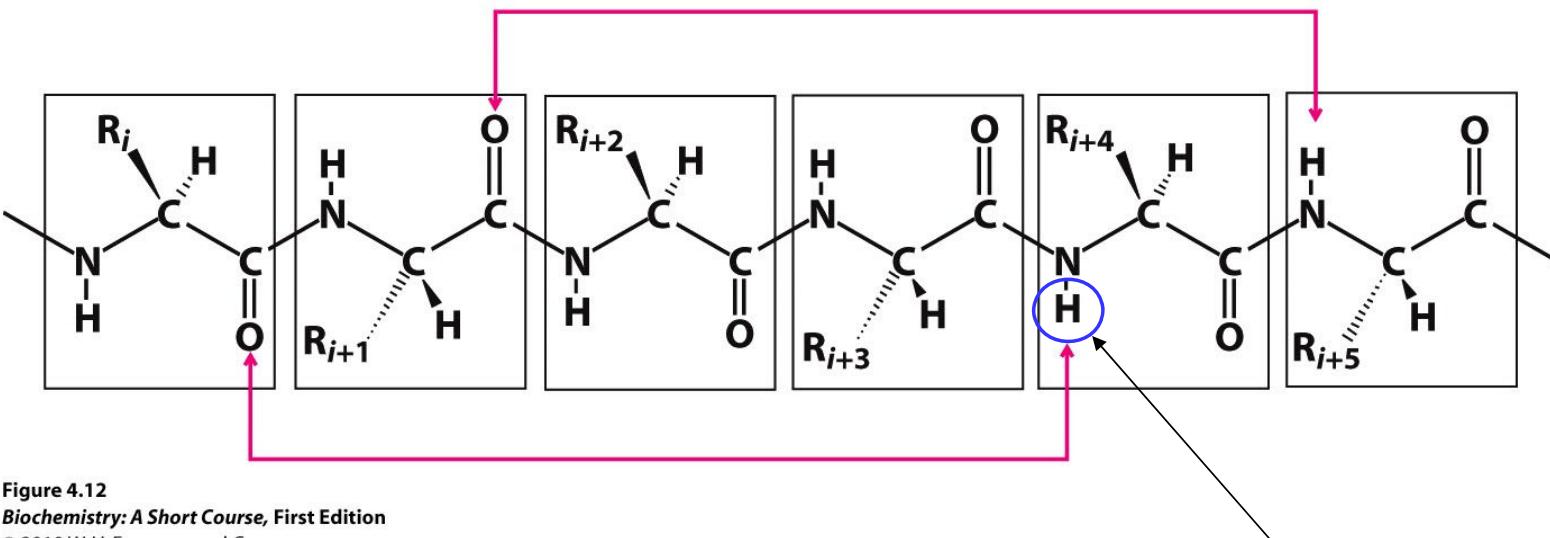


Figure 4.12
Biochemistry: A Short Course, First Edition
© 2010 W.H. Freeman and Company

Proline residues do not have this free H atom, because of the ring structure of its side chain bonding back to the α -N – they therefore break α -helices!

Hydrogen bond patterns in anti-parallel and parallel β -sheets

- ❖ β -strands:
 - opposite direction – **anti-parallel** or
 - same direction – **parallel**
- Look at protein chain direction: N- to C- for each amino acid

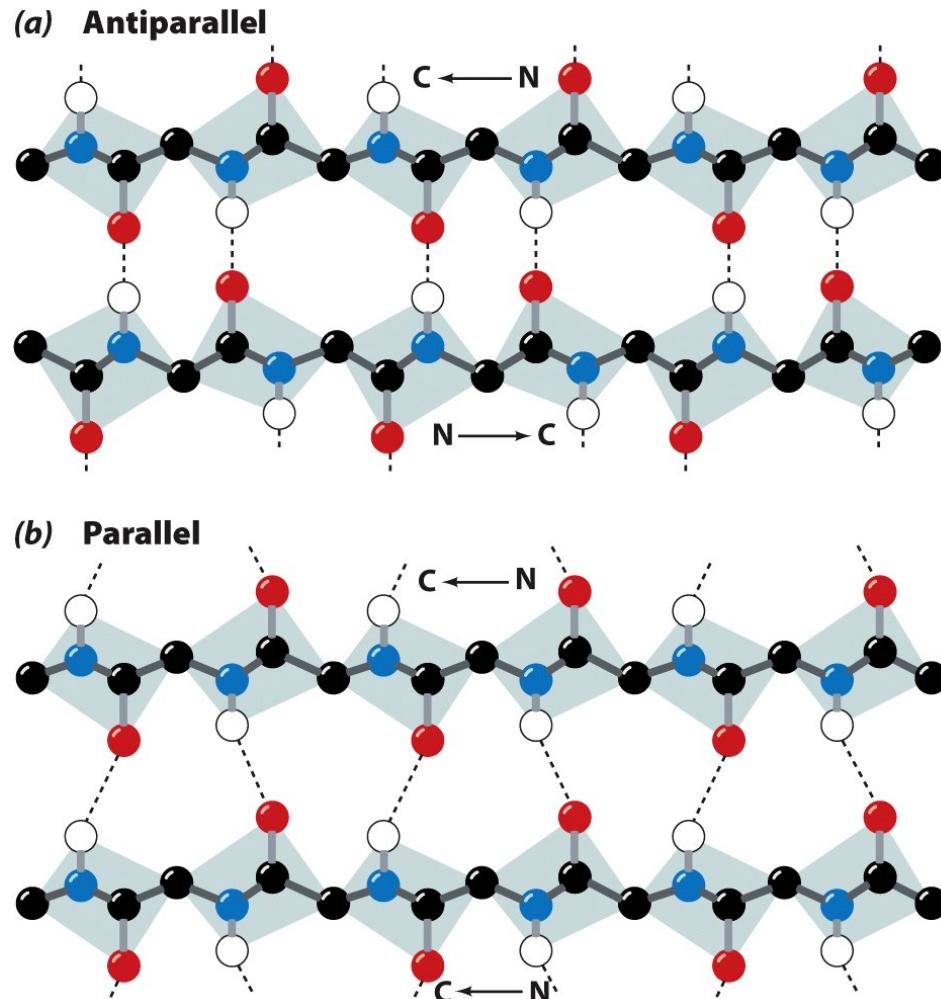


Figure 6-9
Illustration, Irving Geis. Image from the Irving Geis Collection/Howard Hughes Medical Institute. Rights owned by HHMI. Reproduction by permission only.

Connecting Adjacent β Strands by loops or turns

(a)



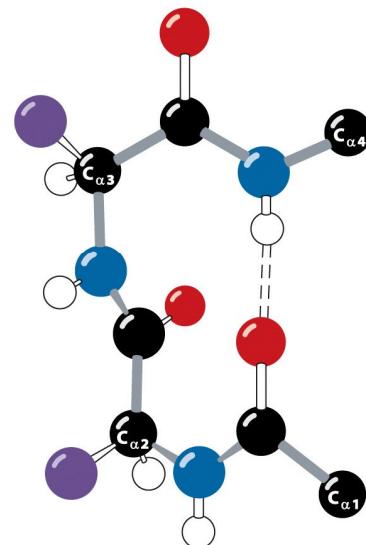
(b)



Figure 6-13
© 2013 John Wiley & Sons, Inc. All rights reserved.

- Turns are short hydrogen bonded segments
- Loops and turns for links between:
 - helices;
 - strands; and
 - helices and strands.

(a) Type I



(b) Type II

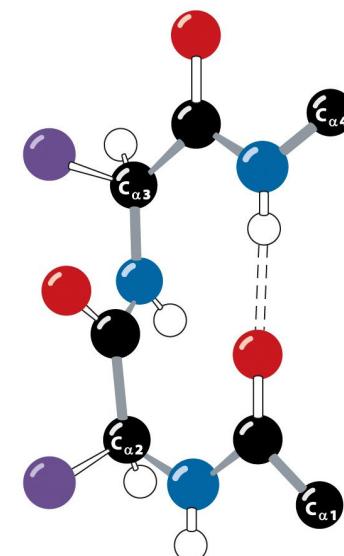


TABLE 6-1 Propensities of Amino Acid Residues for α Helical and β Sheet Conformations

| Residue | P_α | P_β |
|---------|------------|-----------|
| Ala | 1.42 | 0.83 |
| Arg | 0.98 | 0.93 |
| Asn | 0.67 | 0.89 |
| Asp | 1.01 | 0.54 |
| Cys | 0.70 | 1.19 |
| Gln | 1.11 | 1.10 |
| Glu | 1.51 | 0.37 |
| Gly | 0.57 | 0.75 |
| His | 1.00 | 0.87 |
| Ile | 1.08 | 1.60 |
| Leu | 1.21 | 1.30 |
| Lys | 1.16 | 0.74 |
| Met | 1.45 | 1.05 |
| Phe | 1.13 | 1.38 |
| Pro | 0.57 | 0.55 |
| Ser | 0.77 | 0.75 |
| Thr | 0.83 | 1.19 |
| Trp | 1.08 | 1.37 |
| Tyr | 0.69 | 1.47 |
| Val | 1.06 | 1.70 |

Source: Chou, P.Y. and Fasman, G.D., *Annu. Rev. Biochem.* 47, 258 (1978).

Sequence Affects 2° Structure

Unhappy in α -helix:
Pro, Gly, Asn, Tyr, Ser, Cys

Unhappy in β -sheet:
Glu, Asp, Pro

Gly prefers loops or turns

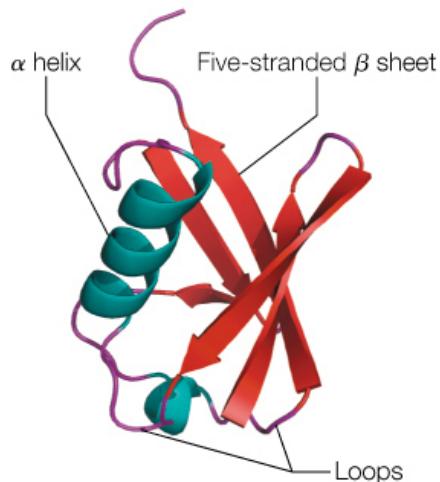
Thus, the aa composition and primary structure of a protein determine its secondary structure.

Protein Secondary Structure - summary

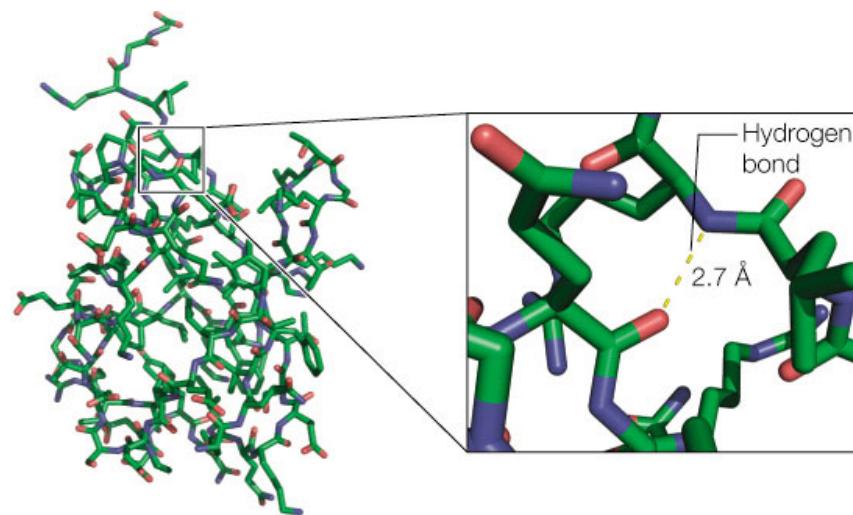
- The planar nature of the peptide group limits the conformational flexibility of the polypeptide chain.
- The α -helix and the β -sheet allow the polypeptide chain to adopt favorable ϕ and ψ angles and to form hydrogen bonds.
- Fibrous proteins contain long stretches of regular secondary structure, such as the “coiled coils” (α -helices twisted together) in α -keratin and the polyproline triple helix in collagen.
- Not all polypeptide segments form regular secondary structure such as α helices or β sheets.

Tertiary structure of globular proteins

Representations of 3D Structures



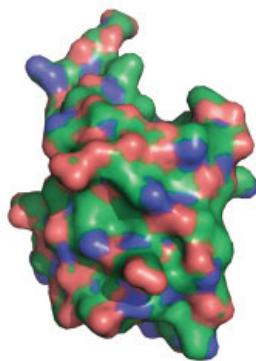
(a) A cartoon model of the protein backbone. An α helix (cyan) is packed against a five-stranded β sheet (red) composed of parallel and antiparallel strands. Loops are shown in magenta.



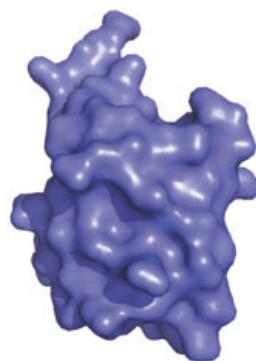
(b) A stick model showing the locations of all atoms (excluding H atoms). C atoms are green, N atoms are blue, and O atoms are red. The inset shows a hydrogen bond of 2.7 Å between main-chain atoms.

FIGURE 6.15 The structure of human ubiquitin.

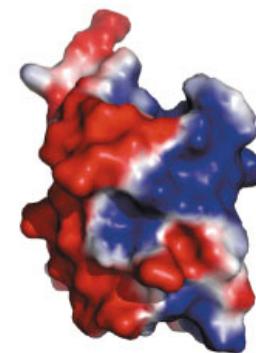
Representations of 3D Structures (2 of 2)



Atom coloring is the same as in panel (b).



Monochrome to emphasize the irregularities of the protein surface.



Distribution of positive (blue) and negative (red) charge density on the protein surface at pH = 7.

(c) Three surface models, showing the solvent-accessible surface of the molecule.

FIGURE 6.15 The structure of human ubiquitin.

Methods used to determine protein structure

- Electron microscopy (see collagen image – slide 18) – for large assemblies and membrane-embedded proteins
- X-ray diffraction of protein crystals (aka X-ray crystallography) – for globular proteins
- Nuclear magnetic resonance (NMR) spectroscopy – for small proteins
- Experimentally solved structures are in the Protein Data Bank (PDB)
- www.rcsb.org

X-Ray Diffraction

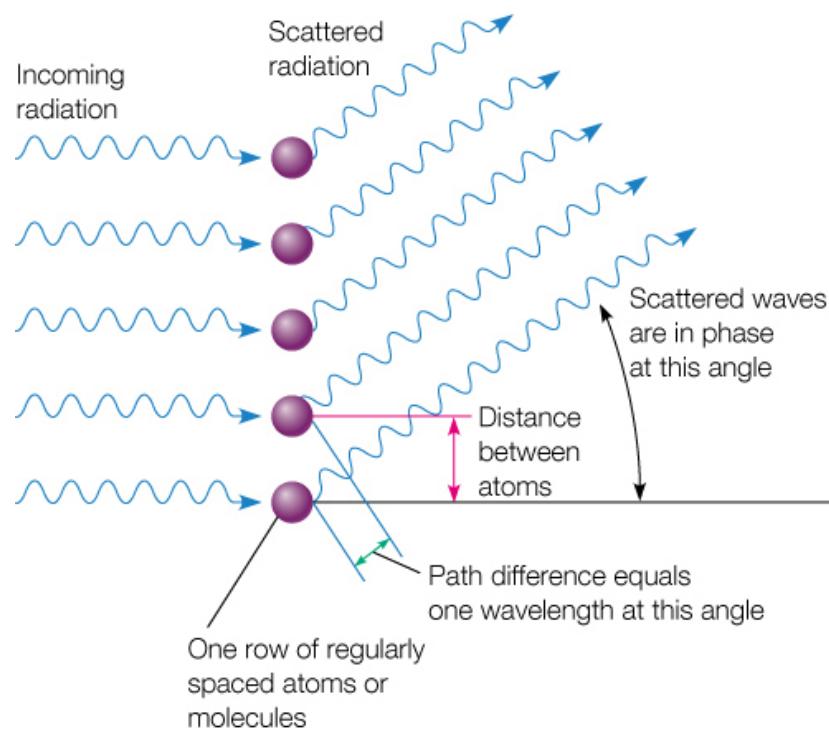


FIGURE 4B.1 Diffraction from a very simple structure—a row of atoms or molecules.

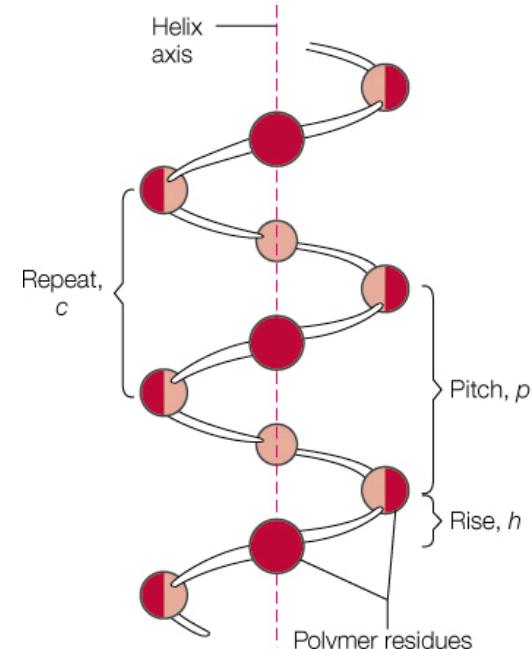


FIGURE 4B.2 A simple helical molecule.

X-ray Diffraction pattern

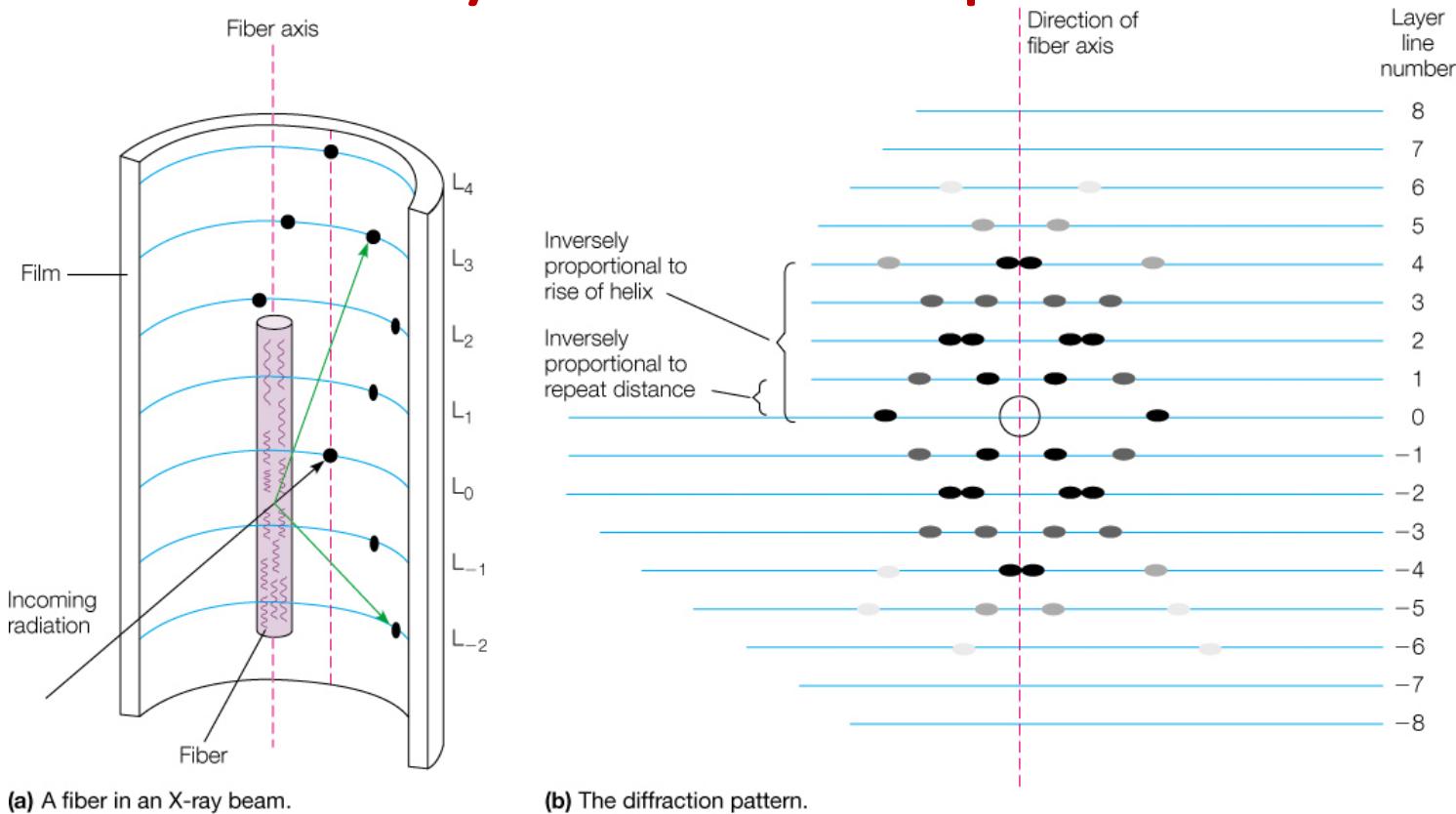
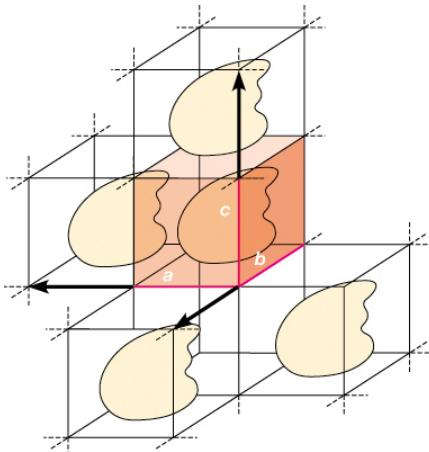


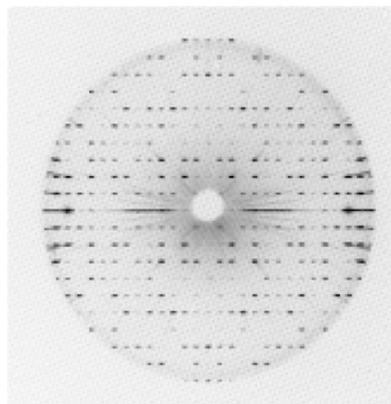
FIGURE 4B.3 Diffraction from fibers.



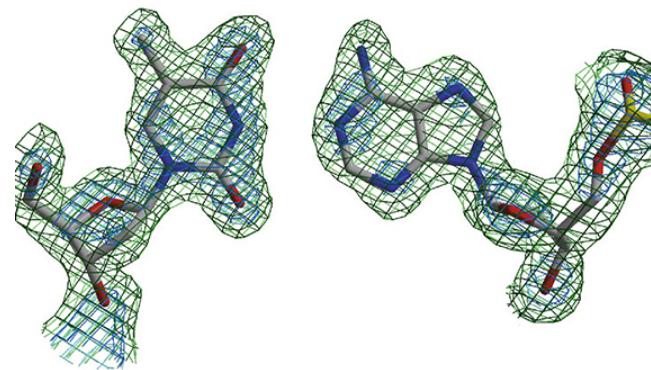
Crystal Diffraction



Molecules in a unit cell



Diffraction pattern of a small DNA crystal



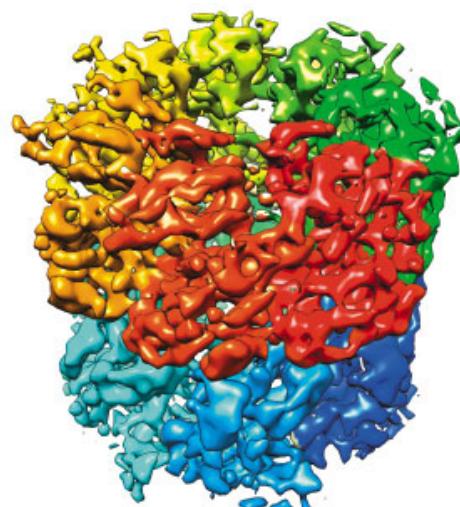
Partial electron density map derived from the diffraction pattern

Crystal Diffraction

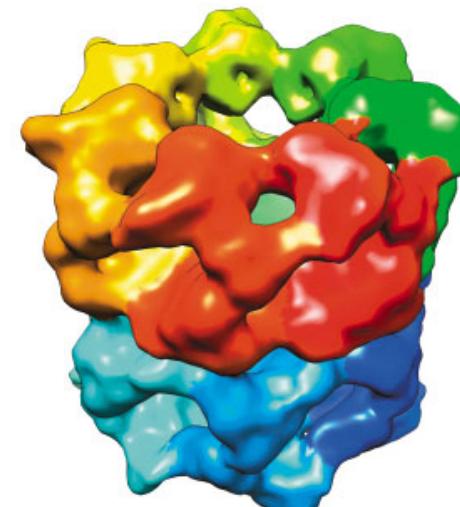
A model structure (GroEL) at different resolutions



4 Å



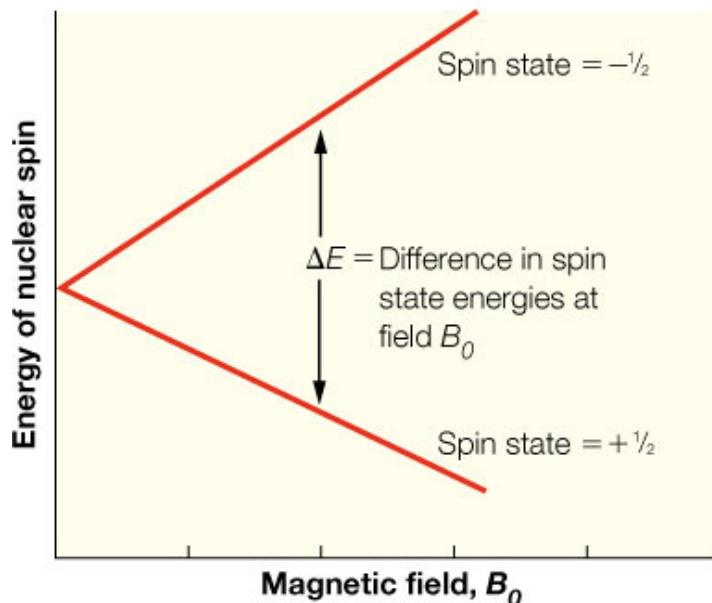
8 Å



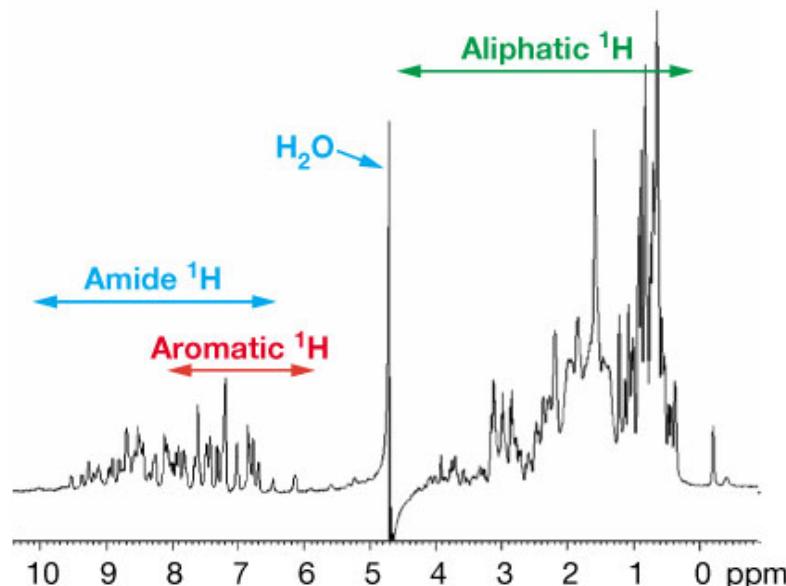
16 Å



NMR Spectroscopy



(a) The effect of magnetic field strength on the energies of nuclear spin states (e.g., ^1H , ^{13}C).



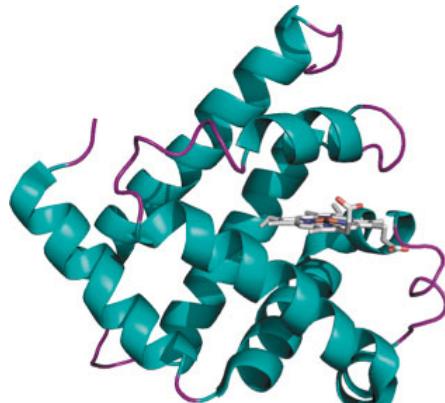
(b) A 500 MHz ^1H NMR spectrum of human ubiquitin (1 mM ubiquitin in 25 mM sodium phosphate, 150 mM NaCl, pH 7.0). This protein has 76 residues, which give rise to ~600 peaks in the ^1H NMR spectrum. The x-axis is the chemical shift, δ in parts per million (ppm).

Figure 6A.9, Nuclear magnetic resonance spectroscopy.

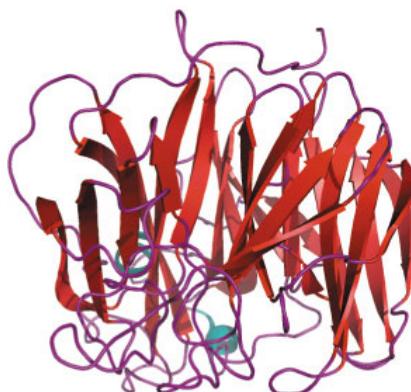
Folding into Defined Structures with Diverse Functions

- Proteins have diverse structures, with varying amounts of helix, sheet, and loop regions
- Larger proteins often contain two or more distinct “domains” of compact folded structure
- A typical protein “domain” is ~200 amino acids and will fold independently
- A domain frequently possesses some defined function (e.g., DNA recognition, oligomerization, cofactor binding, etc.)

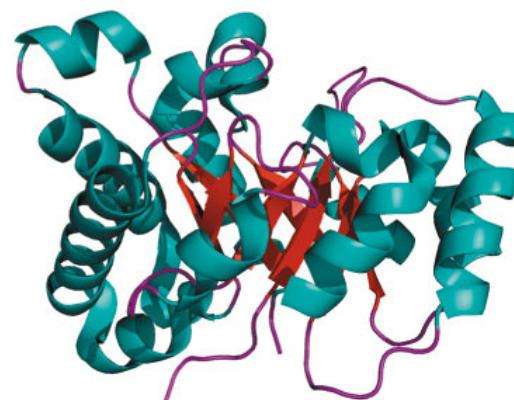
Folding into Defined Structures with Diverse Functions



Myoglobin



Neuraminidase



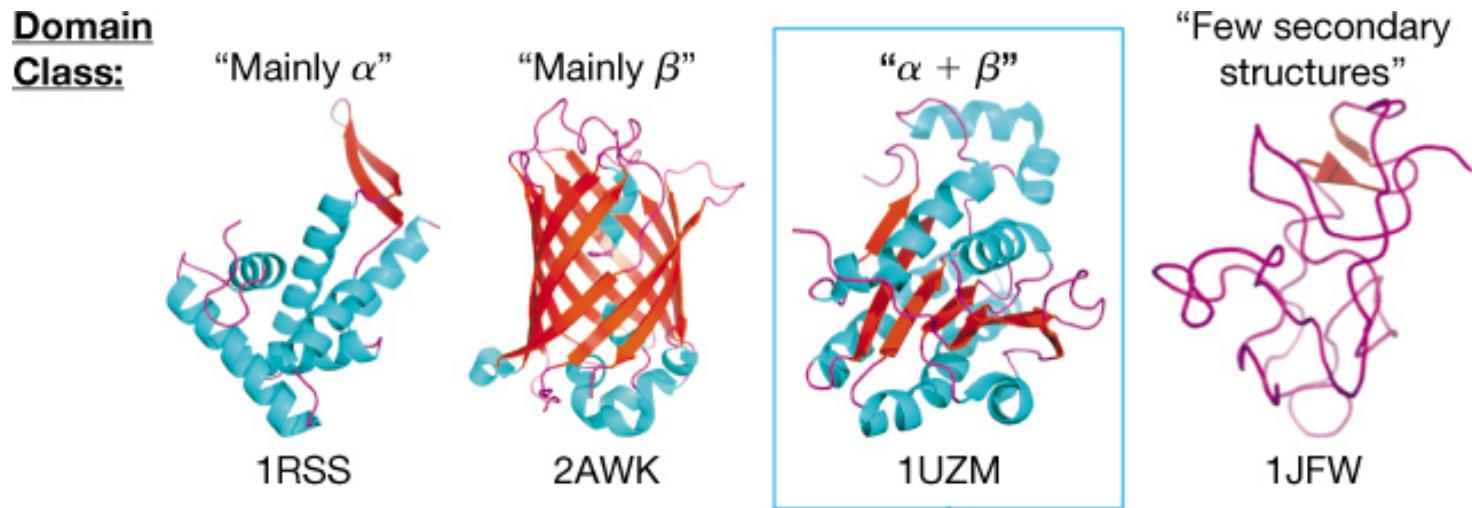
Triosephosphate isomerase (TIM)



Classification of Protein Structure

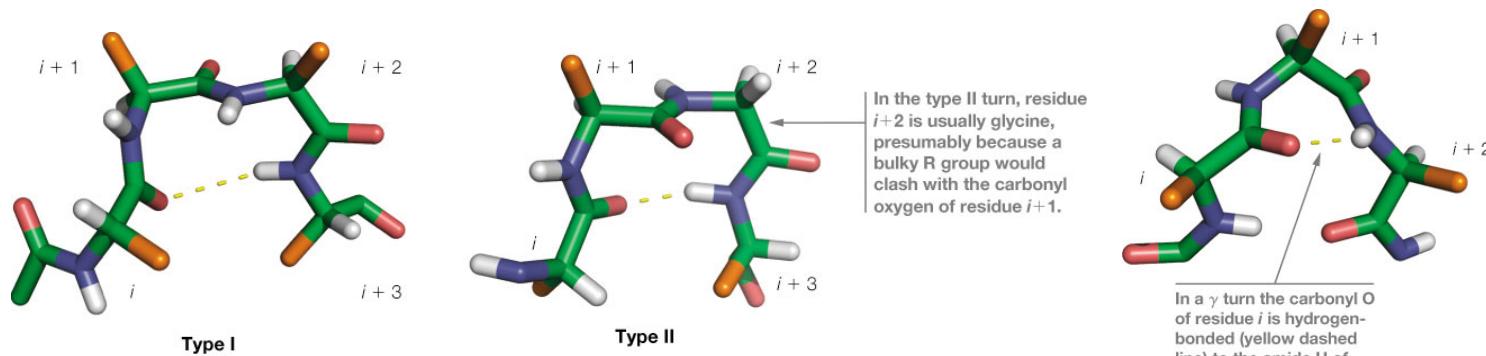
- Protein tertiary structure is characterized by the content of helix and sheet secondary structures, and by defined turns that link these secondary structures
- Some proteins are predominantly helix or sheet, others possess a mixture of helix and sheet, or very little defined secondary structure
- Not all parts of a globular protein structure can be categorized as helix, sheet, or turn. Such regions are often called “random coil” or, more properly, “irregularly structured regions.”

Classification of Protein Structure by Domains



Common Features of Folded Globular Proteins

- Globular proteins have a nonpolar (hydrophobic) interior and a more hydrophilic exterior
- β -sheets are usually twisted or wrapped into barrel structures
- The polypeptide chain can turn corners, for example, β -turns (type I and II left) or γ -turns (right)



Common Features of Folded Globular Proteins

Distribution of hydrophobic and hydrophilic residues in myoglobin

VLSEGEWQLV LHWWAKVEAD VAGHGQDILR LFLKSHPTEL EKFDRFKHLK
TEAEMKASED LKKHGVTVLT ALGAILKKKG HHEAEALKPLA QSHATKHKIP
IKYLEFISEA IIHVVLHSRHP GDFGADAGQA MNKALELFRK DIAAKYKELG
YQG

Hydrophobic residues are green,
hydrophilic ones are magenta, and
ambivalent ones are black

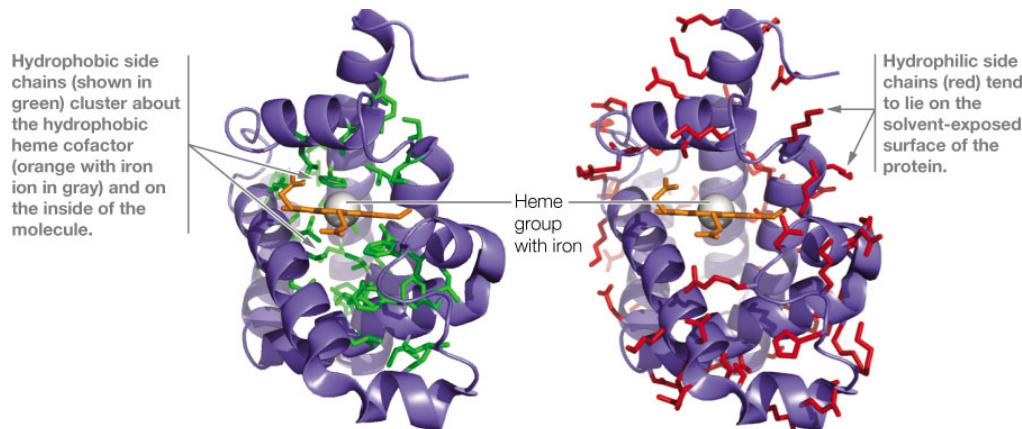


FIGURE 6.19 The distribution of hydrophilic and hydrophobic residues in globular proteins.

Where do the side chains go?

- Based on polarity
 - **Nonpolar** (Val, Leu, Ile, Met, Phe) prefer the **hydrophobic interior**
 - **Charged** polar residues (Asp, Glu, His, Lys, Arg) are usually in contact with aqueous solvent and thus on the **surface**.
 - Uncharged polar groups (Ser, Thr, Asn, Gln, Tyr) are usually on the surface but when buried, form hydrogen bonds.
 - Amino acid mutations may disturb this distribution, and hence affect the stability and the 3D structure.
- **Hydrophobicity** is the **main driving force** for protein tertiary structure.

Thermodynamics, Folding and Stability of Proteins

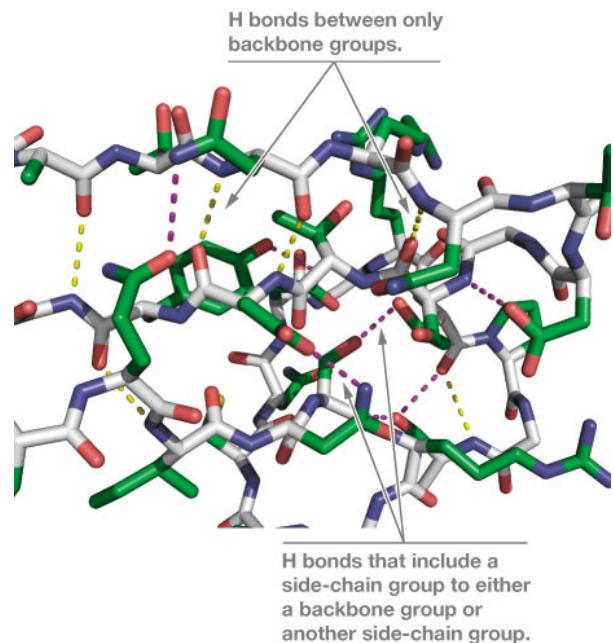


FIGURE 6.23 Details of hydrogen bonding in a typical protein.

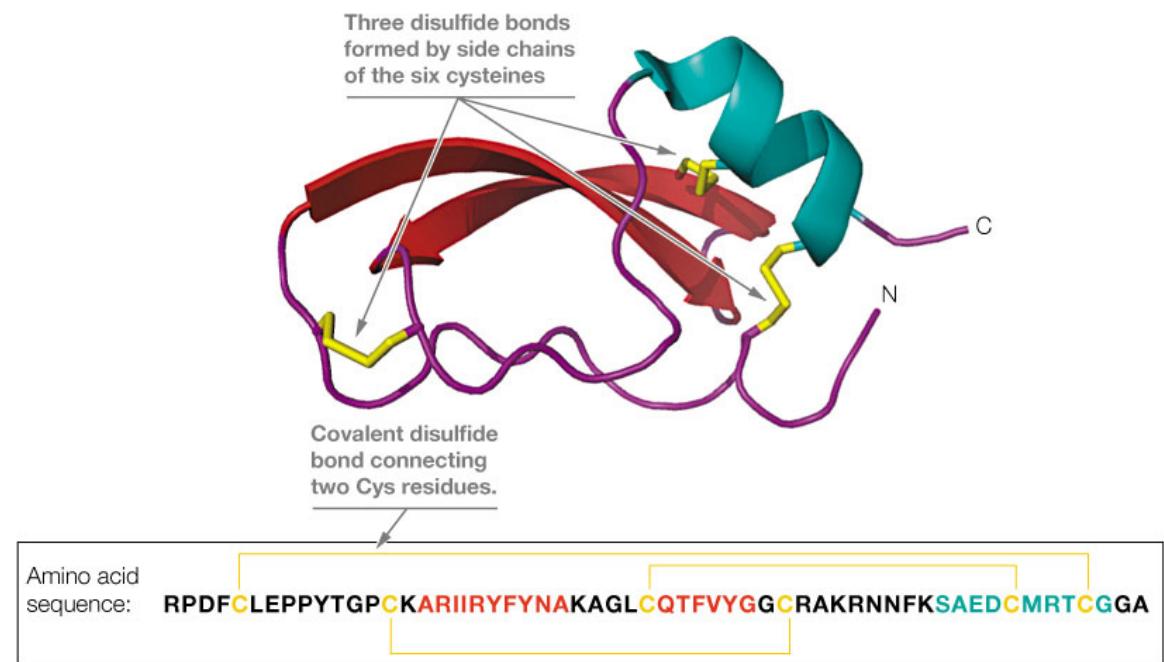
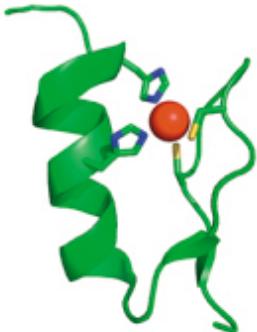


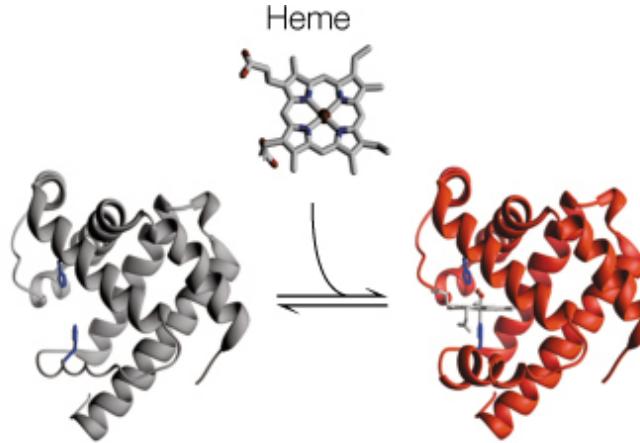
FIGURE 6.25 Disulfide bonds in bovine pancreatic trypsin inhibitor (BPTI).

Thermodynamics, Folding and Stability of Proteins



A “zinc finger” domain bound to a Zn^{2+} ion

(a) The side chains from two histidines and two cysteines bind specifically to a Zn^{2+} ion (red sphere) in a zinc finger domain. The Zn finger domain is a common structure among certain DNA-binding proteins (PDB ID: 1tf6).



Apomyoglobin
(no heme bound)

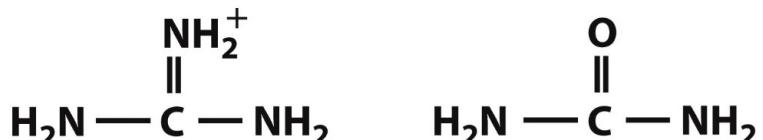
Holomyoglobin
(heme bound)

(b) Heme binding to apomyoglobin stabilizes the folded structure of myoglobin and gives the holoprotein its red color. Holomyoglobin includes both the myoglobin protein and the heme prosthetic group.

FIGURE 6.27 Ion or prosthetic group binding increases protein stability.

Proteins can undergo denaturation and renaturation

- Denaturation leads to **loss of function**.
- Proteins can be denatured by:
 - **Heating**: the entire polypeptide unfolds or “melts”.
Most proteins melt below 100°C.
 - **pH**: changes charge distributions and hydrogen bonding patterns
 - **Detergents**: stabilize hydrophobic side chains and can invert a folded protein.
 - **Chaotropic agents**



Guanidinium ion

Urea

Unnumbered 6 p159
© 2013 John Wiley & Sons, Inc. All rights reserved.

Denatured (unfolded) proteins can be renatured (folded): Anfinsen 1957

- Ribonuclease A (RNase A), a 124-residue single-chain protein, was **denatured by urea**.
- When the urea was removed, the protein **spontaneously renatured** (i.e. **refolded**)!
- The four disulphide bonds re-formed correctly, giving a functional protein.
- Proteins fold spontaneously into their native confirmations under physiological conditions
- Implies the protein's primary structure determines its 3D structure!

Information for Protein Folding

Denaturation and disulphide bond cleavage by β -mercaptoethanol (BME)

- Oxidation and urea removal: scrambled disulphide bonds
- Urea removal followed by oxidation: functional protein recovered!

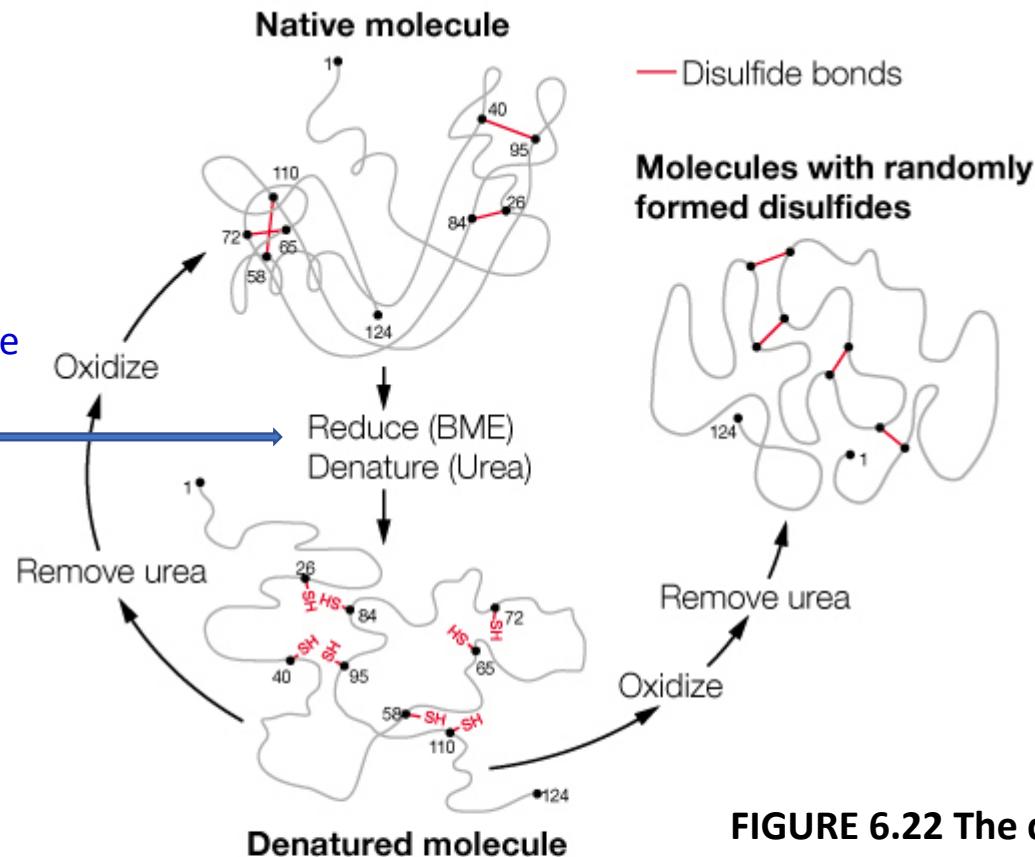
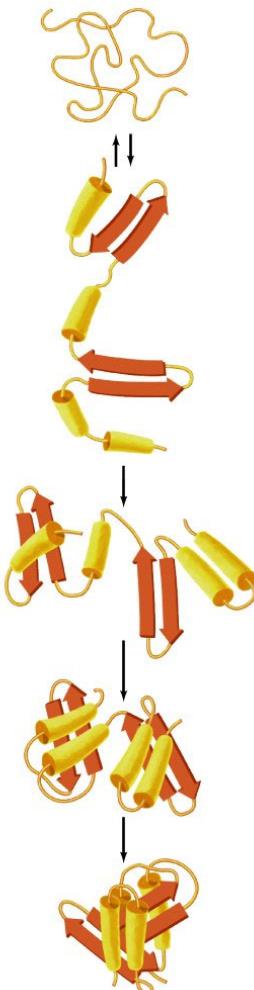


FIGURE 6.22 The denaturation and refolding of ribonuclease A.

- (a) In the classic RNase A refolding experiment of Anfinsen, renaturation before disulfide bond formation yields the active, native conformation, but disulfide bond formation before renaturation yields multiple conformations with little recovery of enzymatic activity.



Hypothetical Protein Folding Pathway



- A protein folds from high energy and high entropy to low energy and low entropy.
- Proteins need to be properly folded for biological activity
- Misfolded proteins are shredded by proteases.
- Proper folding of misfolded proteins is possible with help from large proteins (chaperonins).
- Misfolding can lead to diseases.

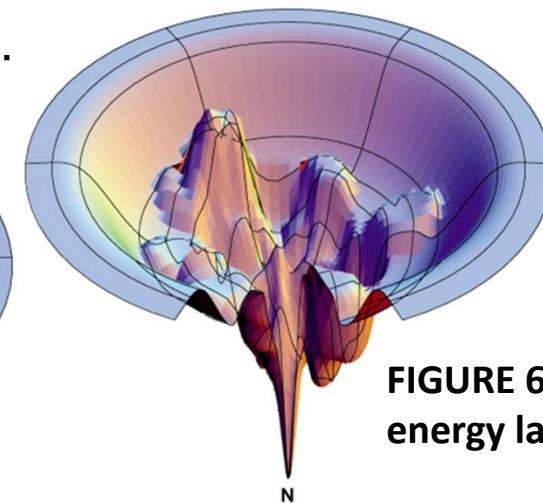
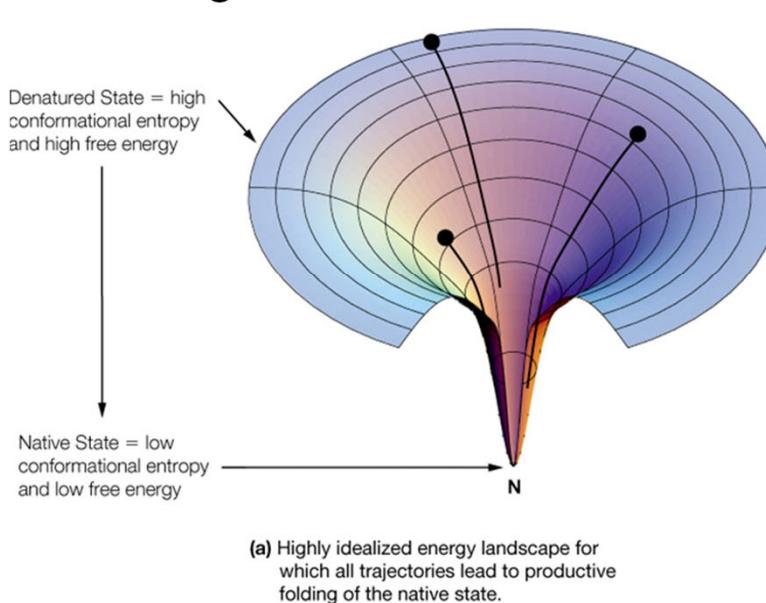


FIGURE 6.29 Protein folding energy landscapes.

(b) A more "rugged" energy landscape. Here, many different paths are possible, some of which lead "downhill" with no local energy minima and give rapid folding. Others may lead to conformations corresponding to local energy minima (i.e., stable intermediate states), which may slow folding (see Figure 6.30).

Models of Protein Folding and Aggregation

Amyloid fibrils and related diseases

Highly ordered amyloids form from non-native folding intermediates or disordered aggregate states

Prions: Infectious agents that cause disease by inducing amyloid formation on contact

TABLE 6.5 Examples of amyloid-related human diseases

| Disease | Associated Protein |
|--|---|
| Alzheimer's disease | Amyloid β or "A β " peptide |
| Parkinson's disease | α -Synuclein |
| Spongiform encephalopathies (such as Creutzfeldt-Jakob disease and kuru) | Prion protein |
| Amyotrophic lateral sclerosis (ALS, or Lou Gehrig's disease) | Superoxide dismutase I |
| Huntington's disease | Huntingtin with polyQ tracts |
| Cataracts | γ -Crystallin |
| Type II diabetes | Islet amyloid polypeptide (IAPP) |
| Injection-localized amyloidosis | Insulin |



Predicting Secondary Structures from Amino Acid Sequences

- Empirical methods are about 80% accurate
- Based on observed distributions of amino acids in helix vs. sheet conformations (e.g., Ala prefers helix, Val prefers sheet, etc.)
- Amphiphilic α -helix shows repeating patterns of side chain polarity every 3–4 residues
- Amphiphilic β strand shows repeating patterns of side chain polarity every other residue

Predicting Tertiary Structures from Amino Acid Sequences

- Critical need for accurate prediction from sequence, since structures are known for only about 1% of all known sequences
- Prediction of tertiary structure is difficult due to the need to correctly predict interactions between residues that are far apart in the primary structure
- Current computational methods are about 60% accurate

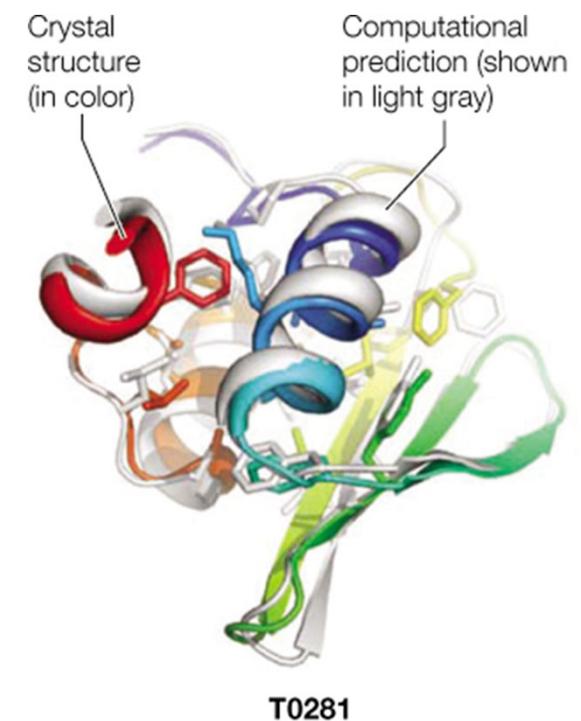
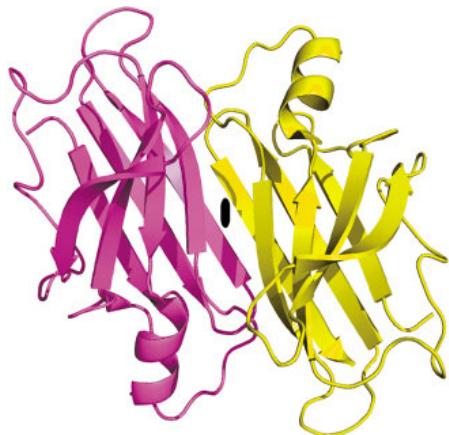
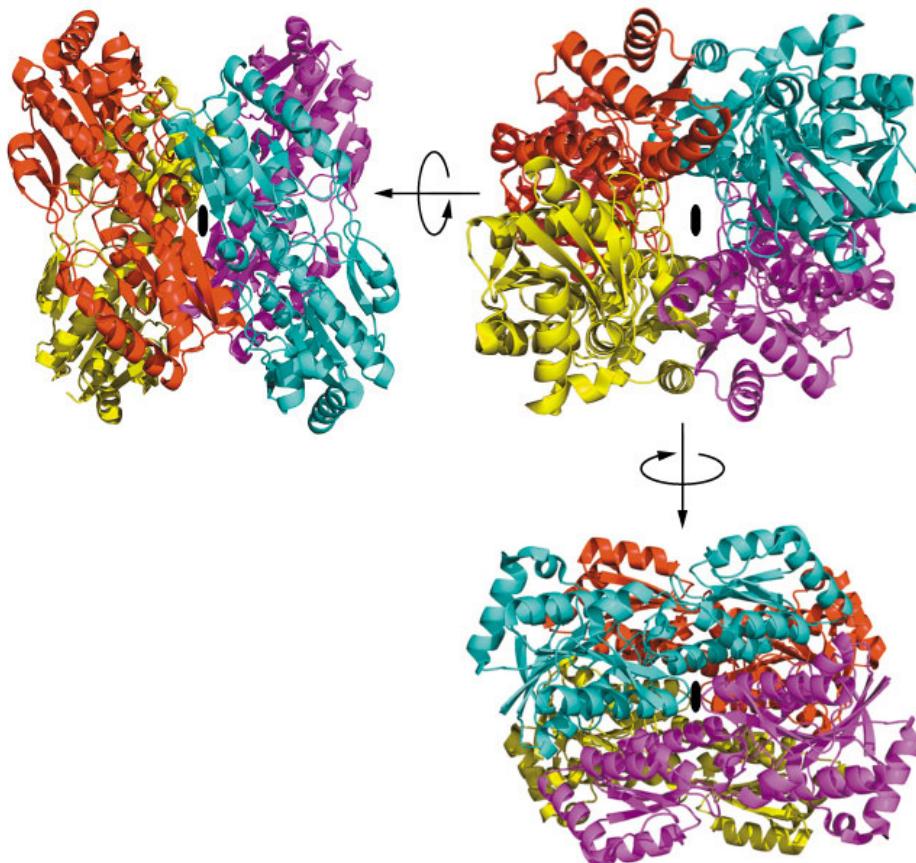


FIGURE 6.33 Comparison of *de novo* prediction to X-ray crystal structures.

Quaternary Structures



(a) In the transthyretin dimer, the two monomers combine to form a complete β sandwich, or flattened β barrel. The dimer has 2-fold symmetry about the C_2 axis perpendicular to the paper (black oval). The isologous interactions are mostly hydrogen bonds between specific β sheet strands.



(b) Three views of the tetrameric enzyme phosphofructokinase. Each view is down one of the three mutually perpendicular C_2 axes.

FIGURE 6.36 Examples of multisubunit proteins.

Quaternary Structures

Interaction between trypsin and the bovine pancreatic trypsin inhibitor (BPTI)

- This is an example of a heterotypic protein–protein interactions (between entirely different proteins)
- Complementary surfaces determine specific interactions

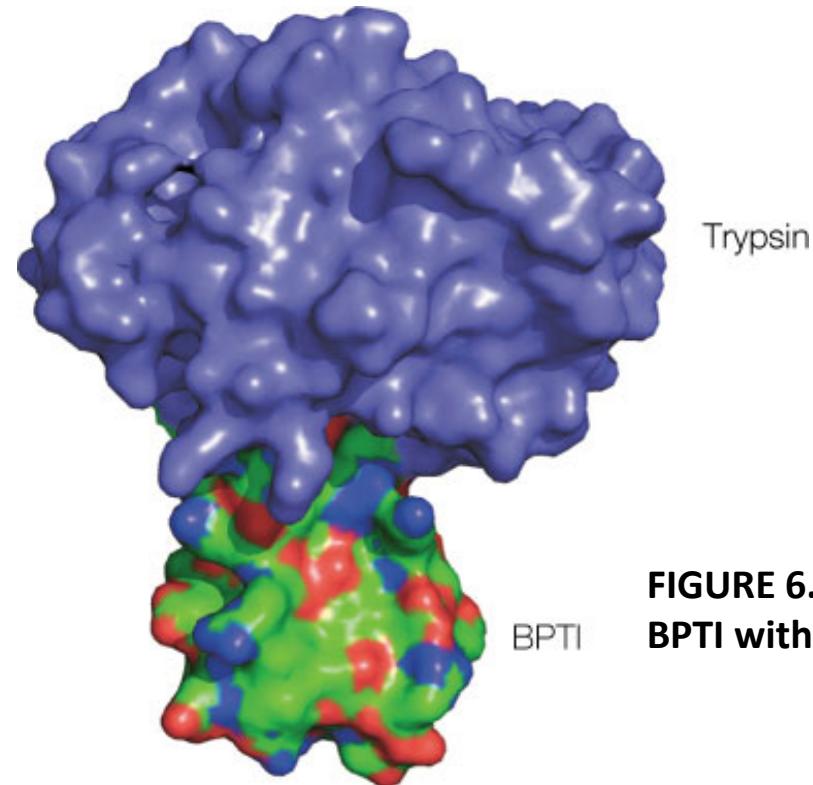


FIGURE 6.37 Interaction of BPTI with trypsin.

Four Levels of Protein Structure

- Primary (1°) structure refers to the amino acid sequence that makes up the protein
- Secondary (2°) structure refers to local areas of repeating main chain structure
- Tertiary (3°) structure refers to the spatial arrangement of the secondary structural elements in the polypeptide chain
- Quaternary (4°) structure refers to the spatial arrangement of multiple polypeptide chains to form multisubunit complexes

Summary

- All proteins have at least three levels of organization
 - Genes dictate the primary level, the amino acid sequences, which then dictates the other levels:
 - secondary,
 - tertiary and, in some cases,
 - quaternary structures.
 - The peptide bond limits the secondary structures that can be formed
 - Ramachandran plots visualize clusters of these sterically allowed secondary structures (α -helix, β -sheet, β_{10} -helix, β -turn, γ -turn, etc.)

Summary

- Proteins can be broadly distinguished into fibrous and globular
- Fibrous proteins are often elongated, performing structural roles in cells, while globular proteins have more complex tertiary structures and fold into compact shapes
- The folding of globular proteins occurs rapidly and spontaneously under standard biochemical conditions
- Many proteins form functional multisubunit assemblies (quaternary structures), for example exhibiting helical symmetry or point-group symmetry

Enzymes 1: Intro and Enzyme Kinetics

Shoba Ranganathan

Applied Biosciences

T: 02 9850 6262; E: shoba.ranganathan@mq.edu.au

Objectives

- What are enzymes?
 - Different types of enzymes
 - Cofactors
- Energetics of enzyme-catalysed reactions
- Free energy as the driving force
 - How to overcome unfavourable energetics
- Enzyme kinetics
 - Experimental determination of kinetic constants.

Textbook Chapter 8



Enzymes: Protein catalysts

A catalyst

- increases the rate or velocity of a chemical reaction without itself being changed in the overall process
- accelerates the approach to equilibrium for a given reaction without changing the thermodynamic favourability of the reaction
- lowers the energy barrier (activation energy of the transition state) in the process of converting a substrate into a product

Enzymes are agents of metabolic change – name from 1878!

- Enzymes cannot catalyse reactions that are energetically unfavourable
- A catalyst is something that reduces the activation energy of a reaction and hence accelerates the reaction rate
 - e.g. Carbonic acid formation
$$\text{CO}_2 + \text{H}_2\text{O} \leftrightarrow \text{H}_2\text{CO}_3$$
 - This reaction proceeds 10 million times faster in the presence of the enzyme (carbonic anhydrase)



Enzyme catalysis

- Reaction rates (i.e. products formed/sec) are 10^6 to 10^{12} times faster with enzymes

TABLE 11-1 Catalytic Power of Some Enzymes

| Enzyme | Nonenzymatic Reaction Rate (s^{-1}) | Enzymatic Reaction Rate (s^{-1}) | Rate Enhancement |
|----------------------------|---|--------------------------------------|----------------------|
| Carbonic anhydrase | 1.3×10^{-1} | 1×10^6 | 7.7×10^6 |
| Chorismate mutase | 2.6×10^{-5} | 50 | 1.9×10^6 |
| Triose phosphate isomerase | 4.3×10^{-6} | 4300 | 1.0×10^9 |
| Carboxypeptidase A | 3.0×10^{-9} | 578 | 1.9×10^{11} |
| AMP nucleosidase | 1.0×10^{-11} | 60 | 6.0×10^{12} |
| Staphylococcal nuclease | 1.7×10^{-13} | 95 | 5.6×10^{14} |

Source: Radzicka, A. and Wolfenden, R., *Science* 267, 91 (1995).

- Milder reaction conditions: $T < 100^\circ C$, $P = 1$ atm and pH almost neutral.
- Greater specificity: in both substrates and products.
- Capacity for regulation: allosteric and covalent modifications and the amount of enzyme synthesized.



Enzymes are classified by the type of reaction they catalyse

- Kinase: transfers a phosphate from ATP to another molecule
- Alcohol dehydrogenase: oxidation of alcohols
- Common name, which comes from the name of the substrate followed by the suffix “-ase”
 - The enzyme, **glucose-6-phosphatase** hydrolyses the phosphate group from glucose-6-phosphate

TABLE 11-2 Enzyme Classification According to Reaction Type

| Classification | Type of Reaction Catalyzed |
|--------------------|--|
| 1. Oxidoreductases | Oxidation-reduction reactions |
| 2. Transferases | Transfer of functional groups |
| 3. Hydrolases | Hydrolysis reactions |
| 4. Lyases | Group elimination to form double bonds |
| 5. Isomerases | Isomerization |
| 6. Ligases | Bond formation coupled with ATP hydrolysis |

Table 11-2
© 2013 John Wiley & Sons, Inc. All rights reserved.



Major Classes of Enzymes

TABLE 8.1 Examples of each of the major classes of enzymes

| Class | Example (reaction type) | Reaction Catalyzed |
|--------------------|--|--|
| 1. Oxidoreductases | Alcohol dehydrogenase (oxidation with NAD ⁺) | $\text{CH}_3\text{CH}_2\text{OH} \xrightarrow{\text{NAD}^+} \text{CH}_3-\overset{\text{O}}{\underset{\text{H}}{\text{C}}} + \text{NADH} + \text{H}^+$ <p style="text-align: center;">Ethanol Acetaldehyde</p> |
| 2. Transferases | Hexokinase (phosphorylation) | <p style="text-align: center;">D-Glucose D-Glucose-6-phosphate</p> |
| 3. Hydrolases | Carboxypeptidase A (peptide bond cleavage) | <p style="text-align: center;">C-terminus of polypeptide Shortened polypeptide C-terminal residue</p> |
| 4. Lyases | Pyruvate decarboxylase (decarboxylation) | $\text{OOC}-\overset{\text{O}}{\underset{\text{H}}{\text{C}}}-\text{CH}_3 + \text{H}^+ \longrightarrow \text{CO}_2 + \text{H}-\overset{\text{O}}{\underset{\text{H}}{\text{C}}}-\text{CH}_3$ <p style="text-align: center;">Pyruvate Acetaldehyde</p> |
| 5. Isomerases | Maleate isomerase (<i>cis-trans</i> isomerization) | <p style="text-align: center;">Maleate Fumarate</p> |
| 6. Ligases | Pyruvate carboxylase (carboxylation) | $\text{OOC}-\overset{\text{O}}{\underset{\text{H}}{\text{C}}}-\text{CH}_3 + \text{CO}_2 \xrightarrow{\text{ATP}} \text{OOC}-\overset{\text{O}}{\underset{\text{H}}{\text{C}}}-\text{CH}_2-\text{COO}^-$ <p style="text-align: center;">Pyruvate Oxaloacetate</p> |



Enzymes are both substrate- and stereo-specific

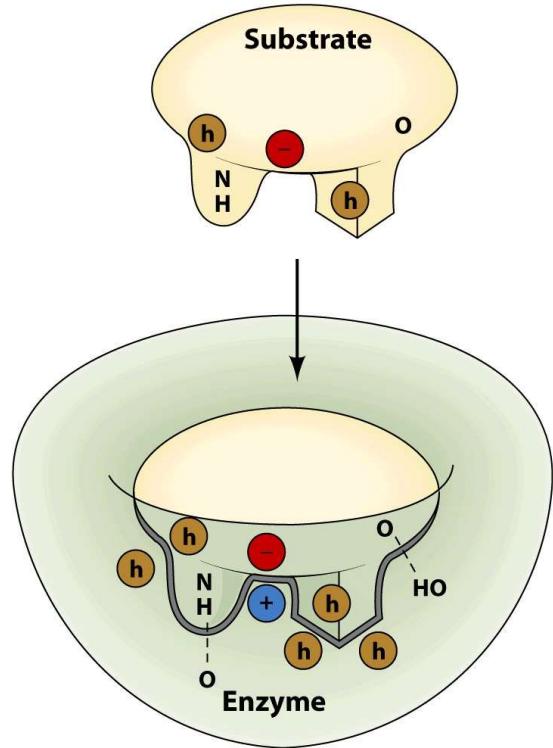


Figure 11-1
© 2013 John Wiley & Sons, Inc. All rights reserved.

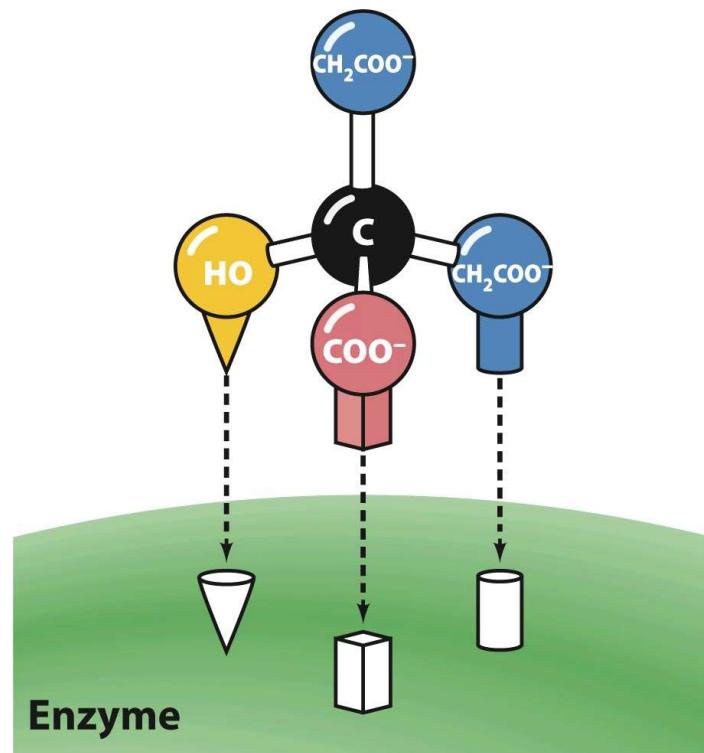
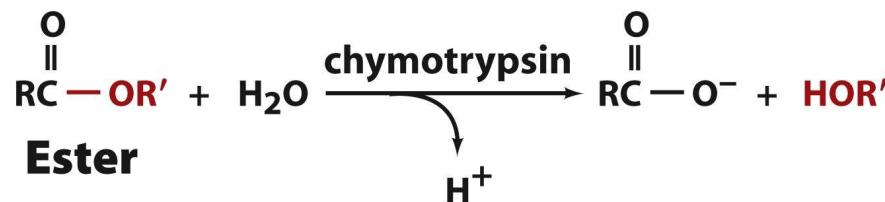
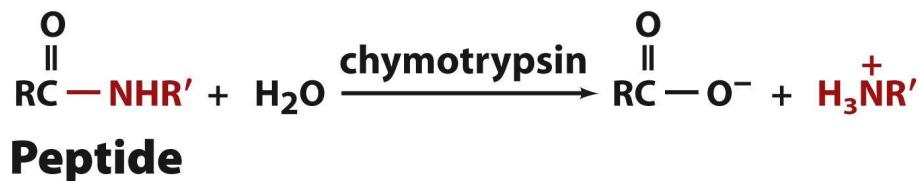


Figure 11-2
© 2013 John Wiley & Sons, Inc. All rights reserved.



Enzymes are substrate specific

- Some enzymes will work on related molecules, although not at maximum efficiency:
 - main role of chymotrypsin is peptide bond hydrolysis
 - Less efficient in hydrolysing synthetic esters!



Unnumbered 11 p318b
© 2013 John Wiley & Sons, Inc. All rights reserved.



Some enzymes require cofactors

- For efficient catalytic activity, enzymes often use **ions or small molecules**
 - These are bound to enzymes, but not irreversibly changed during catalysis.
- These ions or molecules are called coenzymes or cofactors.
- Many vitamins come under this category.

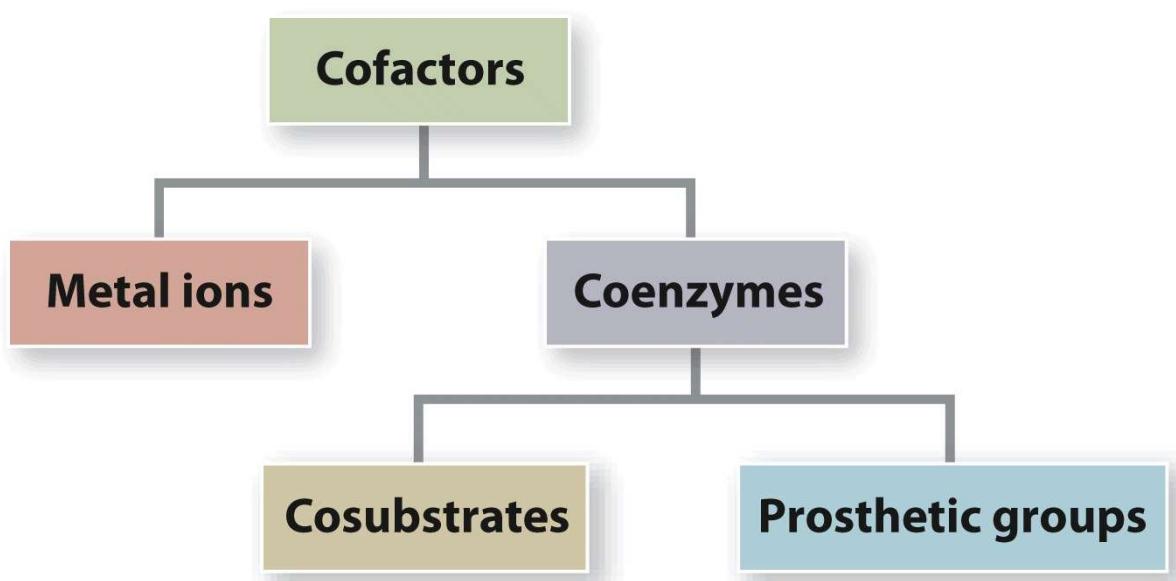


Figure 11-3
© 2013 John Wiley & Sons, Inc. All rights reserved.



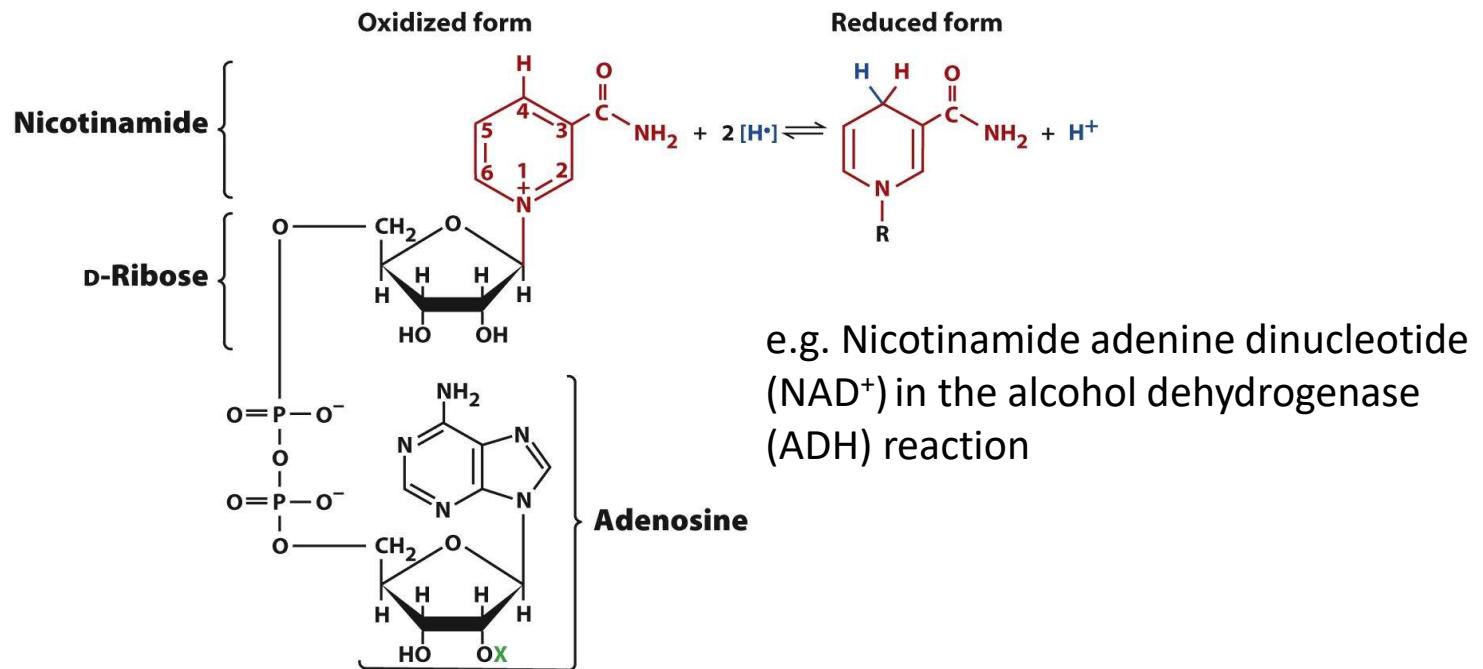
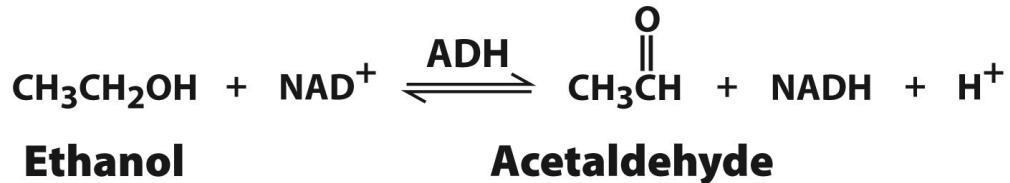
Vitamins provide coenzymes essential for function

TABLE 8.3 Some important coenzymes and related vitamins

| Vitamin | Coenzyme | Reactions involving the coenzyme |
|--|---|--|
| Thiamine (vitamin B ₁) | Thiamine pyrophosphate | Activation and transfer of aldehydes |
| Riboflavin (vitamin B ₂) | Flavin mononucleotide; flavin adenine dinucleotide | Oxidation-reduction |
| Niacin (vitamin B ₃) | Nicotinamide adenine dinucleotide; nicotinamide adenine dinucleotide phosphate | Oxidation-reduction |
| Pantothenic acid (vitamin B ₅) | Coenzyme A | Acyl group activation and transfer |
| Pyridoxine (vitamin B ₆) | Pyridoxal phosphate | Various reactions involving amino acid activation |
| Biotin (vitamin B ₇) | Biotin | CO ₂ activation and transfer |
| Lipoic acid | Lipoamide | Acyl group activation; oxidation-reduction |
| Folic acid (vitamin B ₉) | Tetrahydrofolate | Activation and transfer of single-carbon functional groups |
| Vitamin B ₁₂ | Adenosyl cobalamin; methyl cobalamin | Isomerizations and methyl group transfers |



Some cofactors function as co-substrates



Metal Ions in Enzymes

TABLE 8.4 Metals and trace elements important as enzymatic cofactors

| Metal | Example of Enzyme | Role of Metal |
|-------|-------------------------|---|
| Fe | Cytochrome oxidase | Oxidation-reduction |
| Cu | Ascorbic acid oxidase | Oxidation-reduction |
| Zn | Alcohol dehydrogenase | Helps bind NAD ⁺ |
| Mn | Histidine ammonia lyase | Aids in catalysis by electron withdrawal |
| Co | Glutamate mutase | Co is part of cobalamin coenzyme |
| Ni | Urease | Catalytic site |
| Mo | Xanthine oxidase | Oxidation-reduction |
| V | Nitrate reductase | Oxidation-reduction |
| Se | Glutathione peroxidase | Replaces S in one cysteine in active site |
| Mg | Many kinases | Helps bind ATP |

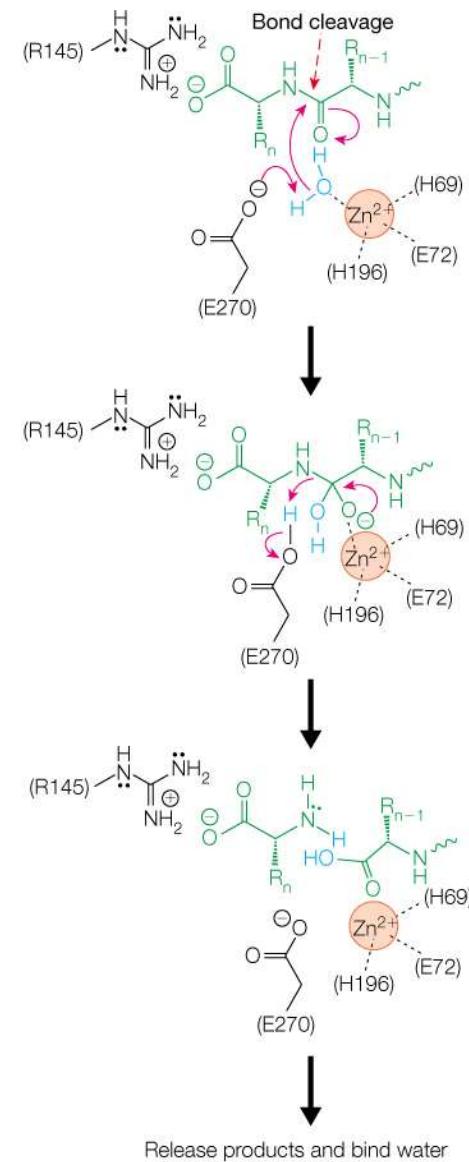


FIGURE 8.19 The mechanism of the protease carboxypeptidase A.

Apo and holo enzymes

- Some enzymes are inactive without the bound cofactor
 - **Apoenzyme (inactive)**
- The active form of the enzyme has the cofactor bound in its specific binding site
 - **Holoenzyme (active)**
- Coenzymes may be chemically changed by the enzymatic reactions they are part of – but they **must be regenerated** for the enzyme activity to continue.



Transition State and Reaction Rate

- The **rate of a chemical reaction** depends on at least one of the following:
 - the order of the reaction;
 - the concentrations of the reactants and products;
 - the temperature;
 - the value of the rate constant
- Whether a chemical reaction is favorable depends on the **free energy difference (ΔG)** between the initial state and the final state (must be negative to occur)

- G , the **Gibbs free energy** (or just “free energy”) is related to the enthalpy (“heat”): H , the entropy (“disorder”): S and the absolute temperature: T

$$G = H - TS, \text{ and}$$

$$\Delta G = \Delta H - T\Delta S$$

- ΔG represents that portion of total energy change that is available to do useful work at constant temperature (T) and pressure (P)

TABLE 3.2 Free energy rules

| If ΔG is . . . | Free energy is . . . | The process is . . . | |
|------------------------|----------------------|--|------------------------|
| Negative | Available to do work | Thermodynamically favorable (and the reverse process is unfavorable) | An exergonic process |
| Zero | Zero | Reversible; the system is at equilibrium | An equilibrium process |
| Positive | Required to do work | Thermodynamically unfavorable (and the reverse process is favorable) | An endergonic process |

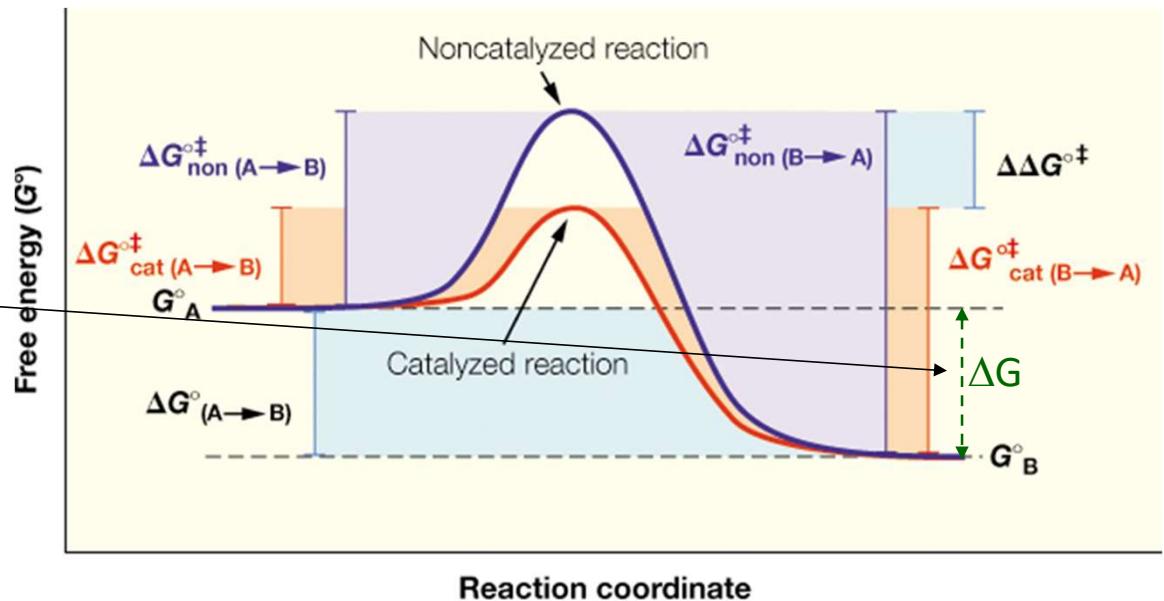
Transition State and Reaction Rate

- The **rate of a chemical reaction** depends on at least one of the following:
 - 1) the order of the reaction;
 - 2) the concentrations of the reactants and products;
 - 3) the temperature;
 - 4) the value of the rate constant
- Whether a chemical reaction is favorable depends on the **free energy difference (ΔG)** between the initial state and the final state (must be negative to occur)
- However, to determine the rate of a chemical reaction, not only its initial state and its final state have to be considered, but also its **transition state**
- The transition state is an additional energy barrier to convert reactants into products (and vice versa)
- The transition state has more free energy in comparison to that of the reactant or product; thus, it is the least stable state during the progress of a chemical reaction



Enzymes lower activation energy of the catalyzed reaction

- Enzymes stabilize **the transition state** and lower the activation energy to achieve rate enhancement
- However, they cannot change the ΔG of the reaction.
 - $\Delta G < 0$: spontaneous reaction
 - $\Delta G > 0$: reverse reaction is spontaneous
- Actual velocity of the reaction depends on kinetics



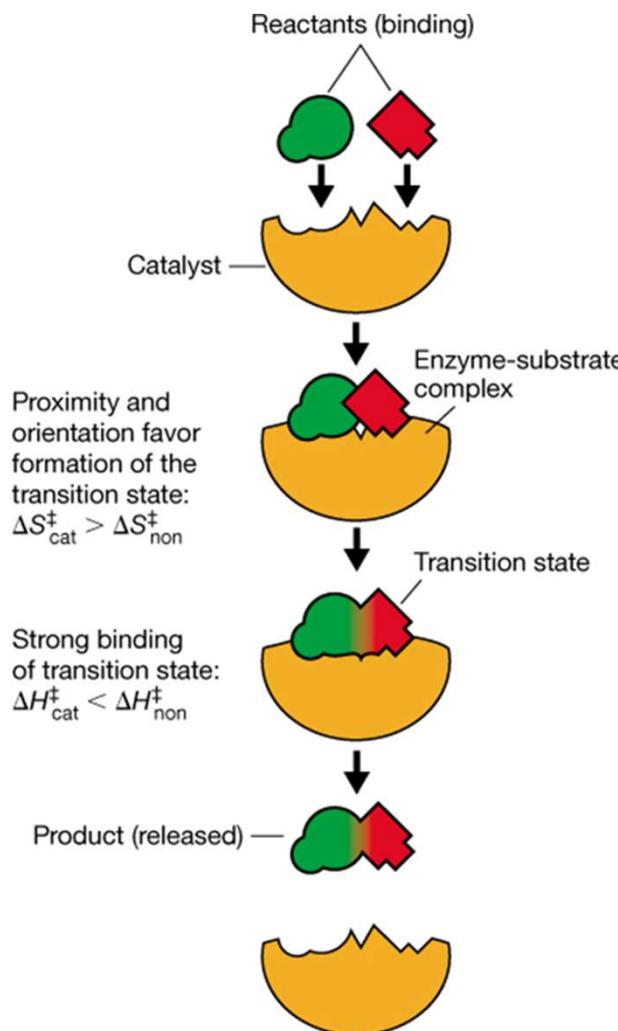


FIGURE 8.6 Entropic and enthalpic factors in catalysis.

Enzymes preferentially bind the transition state (S^{\ddagger}) of the reaction

- The **transition state** of the reaction binds with greater affinity to the enzyme than the substrates or the products.
- **Transition state analogues** are often used as inhibitors in industry and in drug design

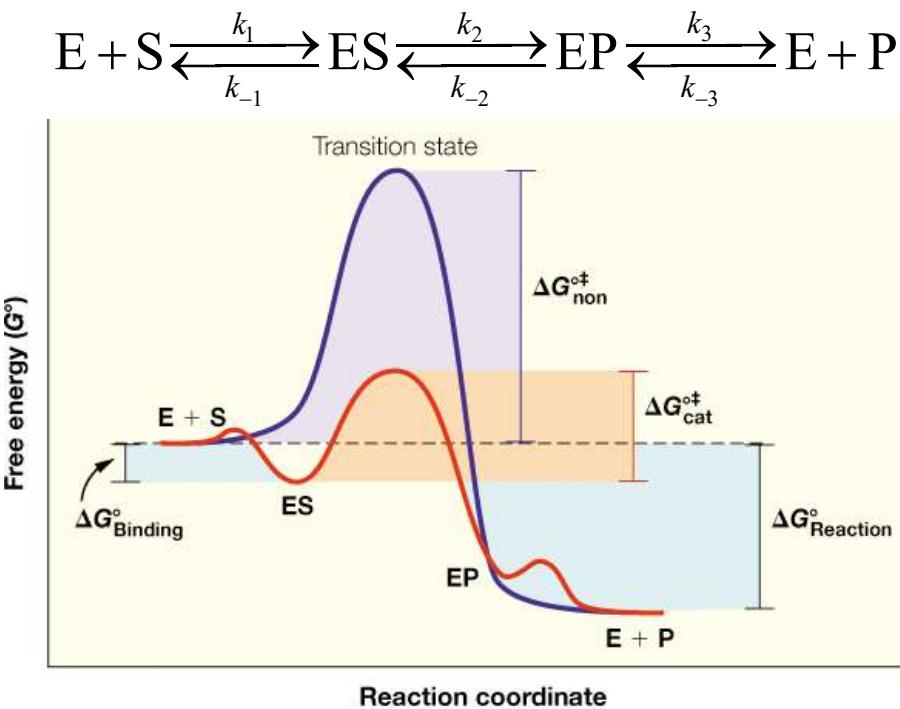
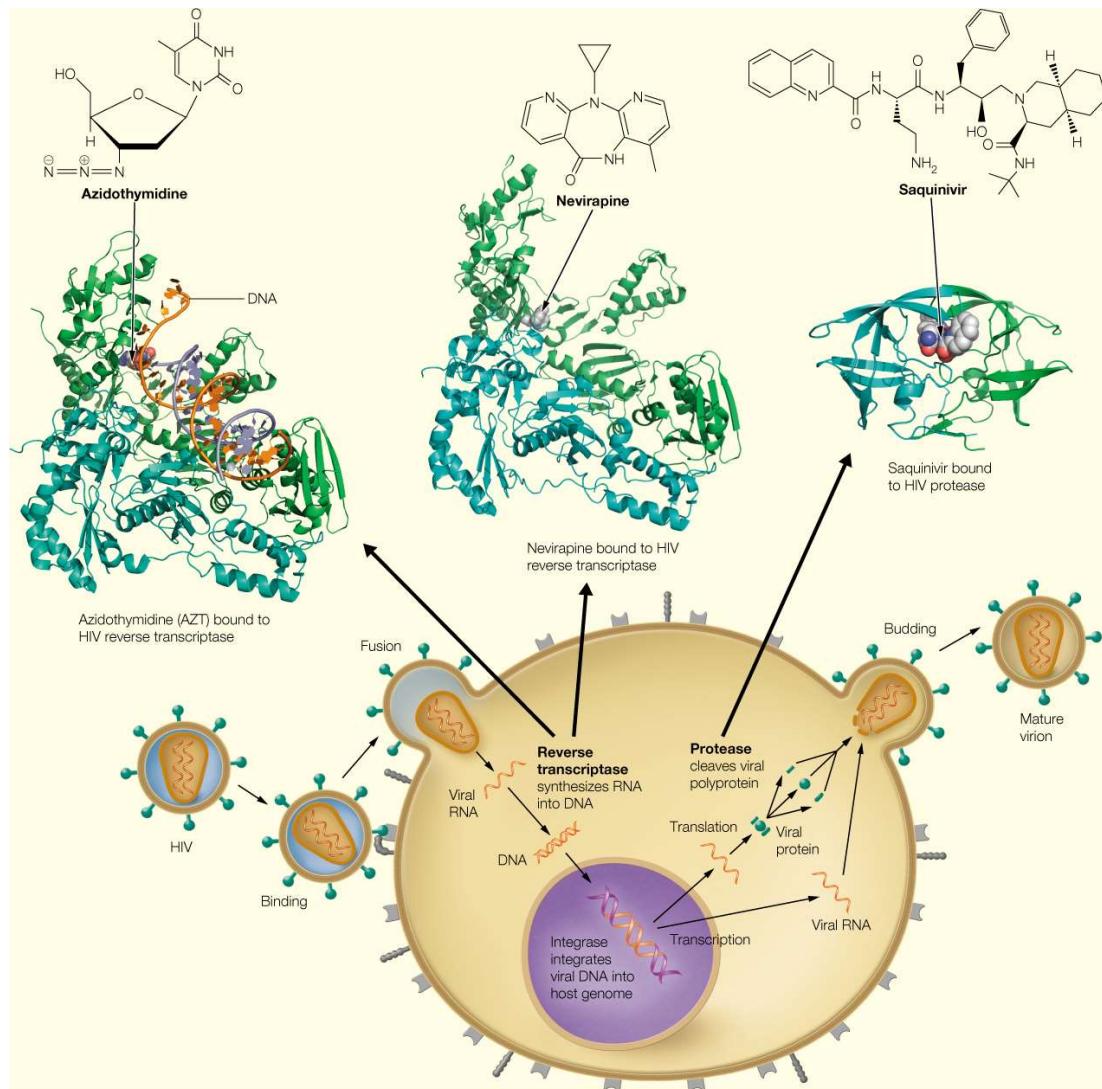


FIGURE 8.10 Reaction coordinate diagram for a simple enzyme-catalyzed reaction.

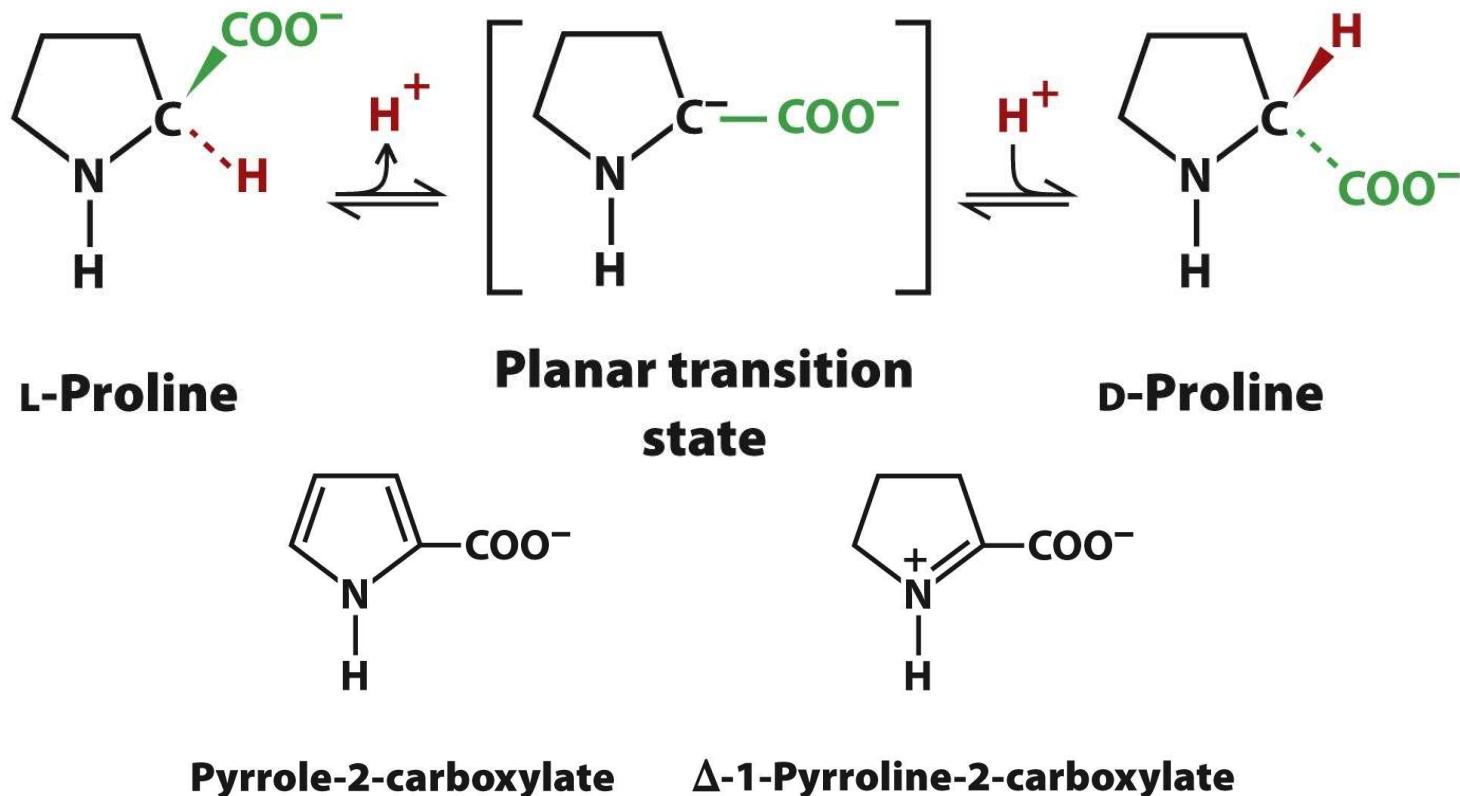
Enzymes and Cell Biology

Enzyme inhibitors as drugs

- Many different therapeutic drugs have been developed to treat HIV infection.
- The three shown here—azidothymidine (AZT), nevirapine, and saquinivir—all have different modes of action against two enzymes critical to the HIV life cycle.
- AZT is a substrate analog that causes premature DNA chain termination.
- Nevirapine distorts the enzyme active site and thereby inhibits reverse transcription.
- Saquinivir blocks the active site of the HIV protease.



Transition state analogs are enzyme inhibitors:
e.g. proline racemase (*Clostridium sticklandii*)



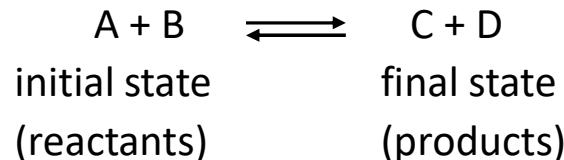
Enzymes Summary

- Enzymes form six classes, based on the reactions they catalyse.
- Enzymes are substrate specific.
- Some enzymes need metal ion cofactors or organic coenzymes (derived from vitamins) to function.
- Enzymes accelerate reactions by lowering activation energy.
 - do **not** change the free energy of the reaction
 - bind preferentially to the **transition state**



Why does a reaction occur?

- A general reaction is:



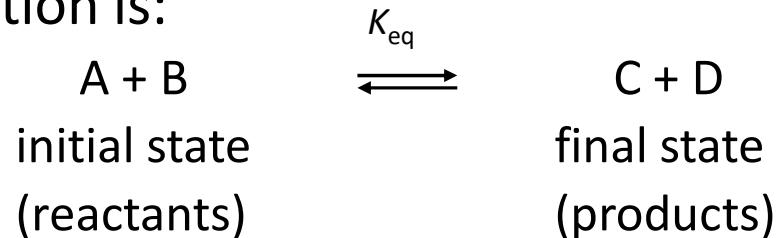
- Free energy of the reaction at constant T and P is:

- $\Delta G = \Delta H - T \Delta S$
 - ΔG is **free energy** change (final state – initial state)
 - ΔH is **enthalpy** change (i.e. heat transfer)
 - T is temperature (in K, Kelvin)
 - ΔS is **entropy** change (a measure of randomness, increasing for spontaneous processes)
- If ΔG is 0, the reaction **is at equilibrium**.
- If ΔG is negative, the reaction **occurs spontaneously**.
- If ΔG is positive, the reaction **does not occur spontaneously**.



Free energy for a reaction

- A general reaction is:



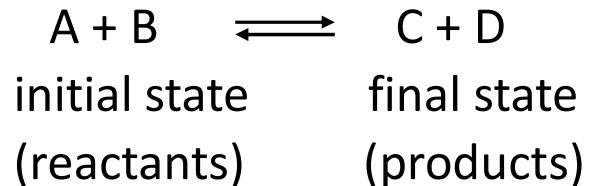
- Standard free energy of the reaction:
 - $\Delta G^\circ = \sum \Delta G^\circ (\text{products}) - \sum \Delta G^\circ (\text{reactants})$
- ΔG versus ΔG°
 - To calculate the free energy change for reactions not at standard state, use

$$\begin{aligned}\Delta G &= \Delta G^\circ + RT \ln ([\text{C}][\text{D}]/[\text{A}][\text{B}]) \\ &= \Delta G^\circ + RT \ln (K_{\text{eq}})\end{aligned}$$



Free energy and equilibrium constant

- For general reaction,
with equilibrium constant K_{eq}



$$\Delta G^\circ = -RT \ln ([C][D]/[A][B])$$

$$\Delta G^\circ = -RT \ln (K_{eq})$$

$$\Delta G^\circ < 2.73; K_{eq} = 10^2$$

- reaction goes to completion

Table 2.5

Relationship between ΔG° and K_{eq} (at 25°C)

| ΔG° (kcal/mole) ^a | K_{eq} |
|---|-----------|
| -6.82 | 10^5 |
| -5.46 | 10^4 |
| -4.09 | 10^3 |
| -2.73 (-11.2 kJ/mole) | 10^2 |
| -1.36 | 10 |
| 0 | 1 |
| 1.36 | 10^{-1} |
| 2.73 | 10^{-2} |
| 4.09 | 10^{-3} |
| 5.46 | 10^{-4} |
| 6.82 | 10^{-5} |



Biological systems perform work

- **Mechanical work**
 - muscle contraction, flagella rotation, chromosome movement
- **Concentration and electrical work**
 - movement of ions and molecules across membranes, osmotic changes, etc.
- **Synthetic work**
 - changes in chemical bonds

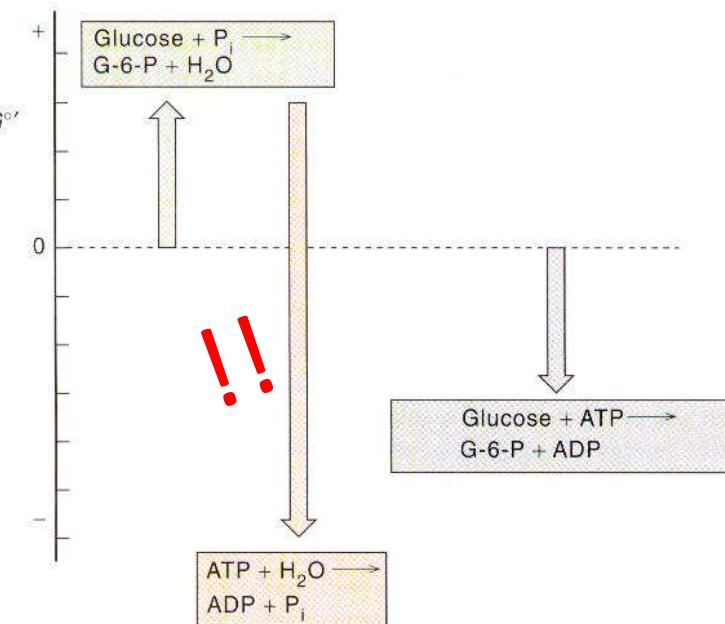
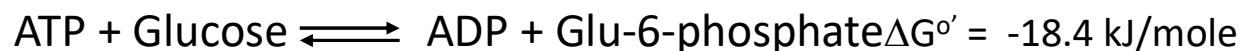
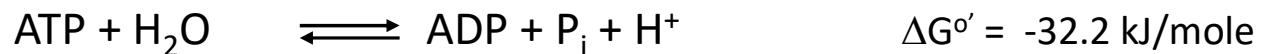
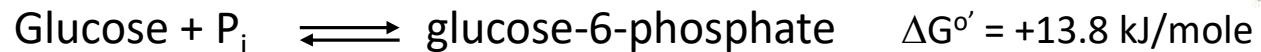
All these actions require **energy**

- Usually from the hydrolysis of high energy phosphate bonds (**ATP, ADP**)
 - Also from **reduced coenzymes**



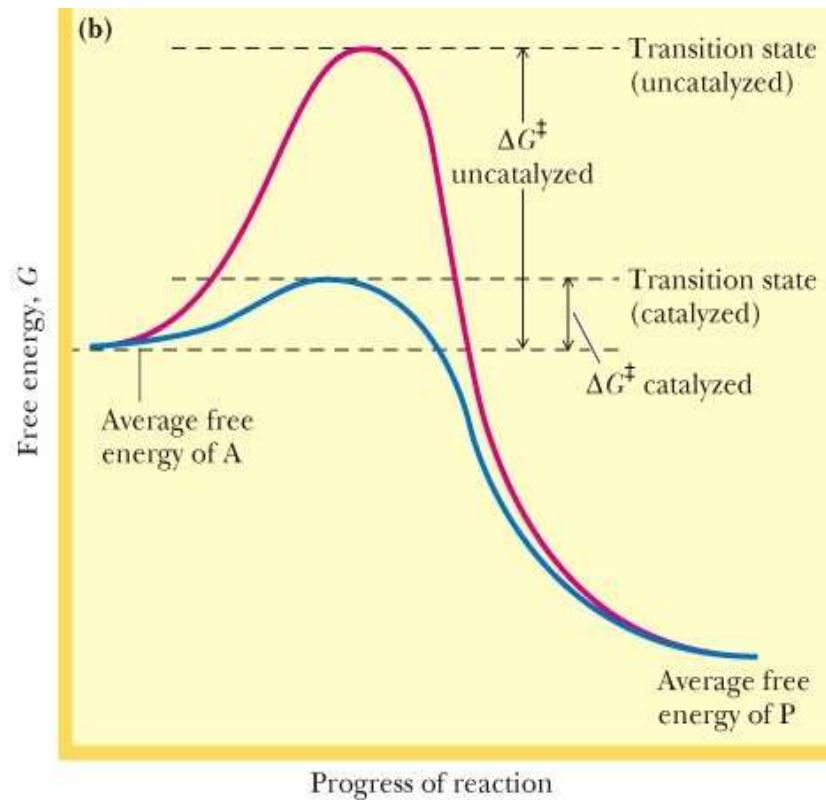
Biological systems can perform energetically unfavorable work

- By coupling an energetically favourable reaction with an unfavourable one:



Enzymes are catalysts

- Enzymes speed up the reaction rate by decreasing the activation energy
 - Uncatalysed:
 $S \leftrightarrow P$
substrate to product
 - Catalysed:
 $E + S \leftrightarrow ES \rightarrow E + P$

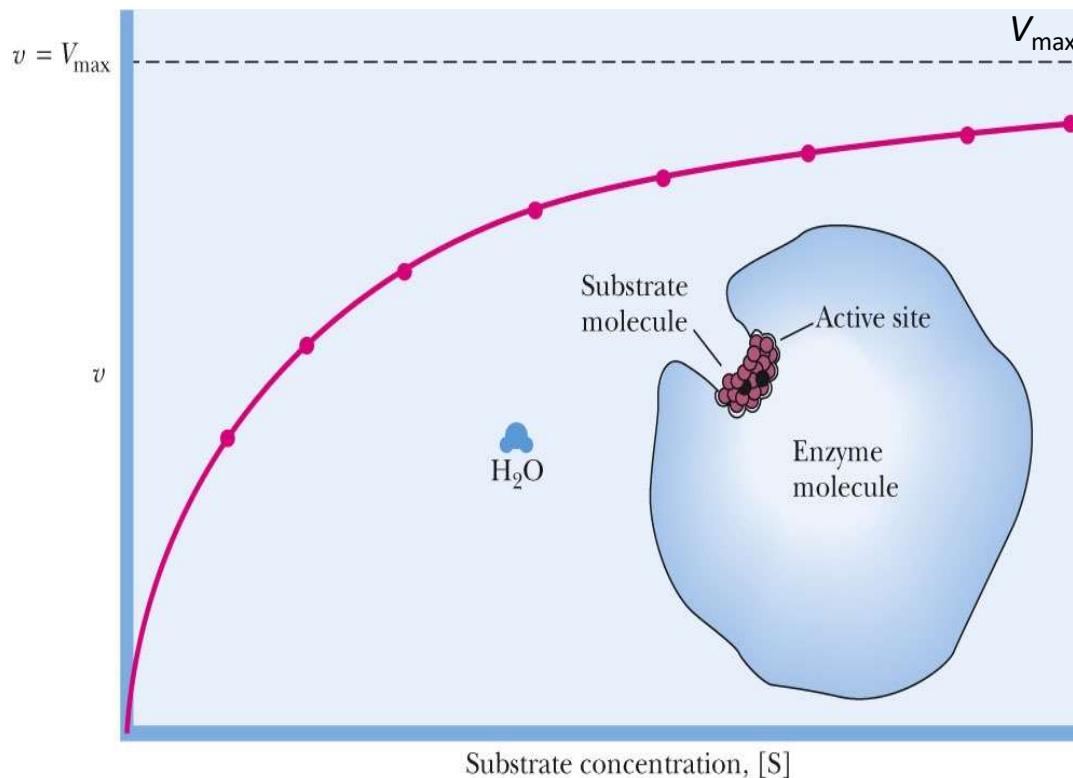


Rate of the reaction

- Substrate S forms product P.
- The rate of the reaction (v , “velocity”) is the rate of P formation.
- v in an uncatalysed reaction is dependent on the starting concentration [S]
 - If we plot the rate vs. the starting concentration of [S] we ideally get a straight line (**known as a first order reaction**)
- For enzyme catalysed, reactions, v is also dependent on the enzyme concentration [E] (**second order reaction**)
 - so if we plot v vs. the starting concentration of [S] we get a hyperbolic curve that reaches a maximum



Rate of the reaction with increasing substrate concentration [S]: Michaelis-Menten kinetics



Concentrations vs. time

- Time-dependent changes in concentrations of S, P, E and ES.
- Only a small quantity of E is needed
- E quickly forms ES, so that $[E] \sim [ES]$
- v is usually measured at constant $[ES]$

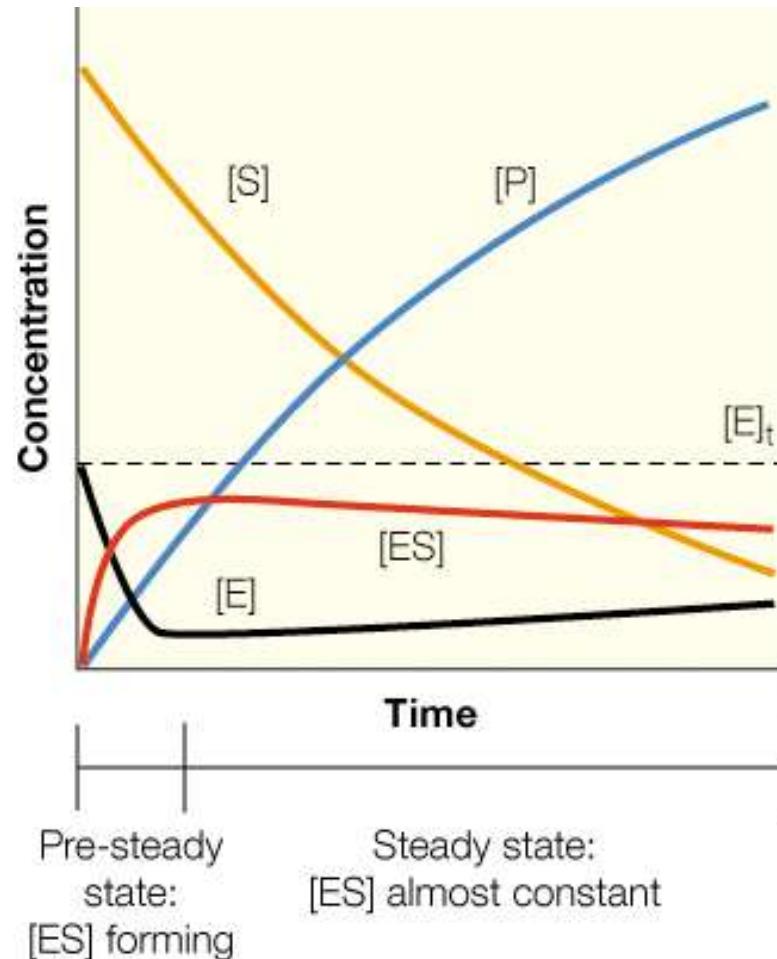


FIGURE 8.20 The steady state in enzyme kinetics.

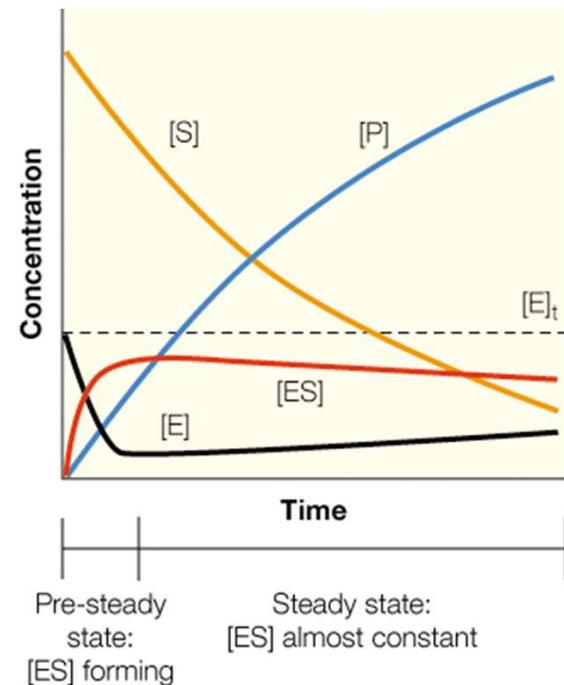
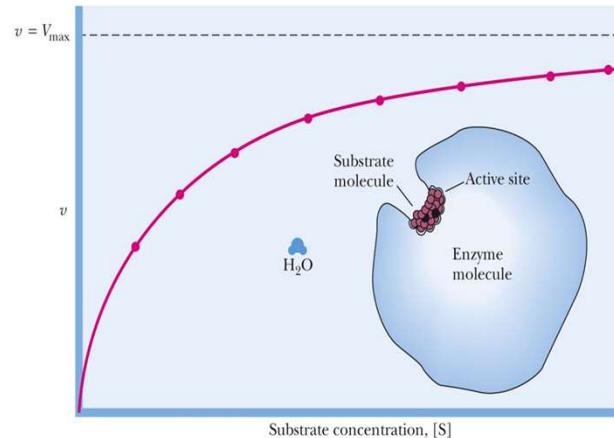


Analysis of enzyme activity – initial velocity and steady state kinetics

- $[ES]$ remains constant for the duration of the steady state
- **The Michaelis-Menten equation (1913!)** describes the hyperbolic graph for v vs. $[S]$.

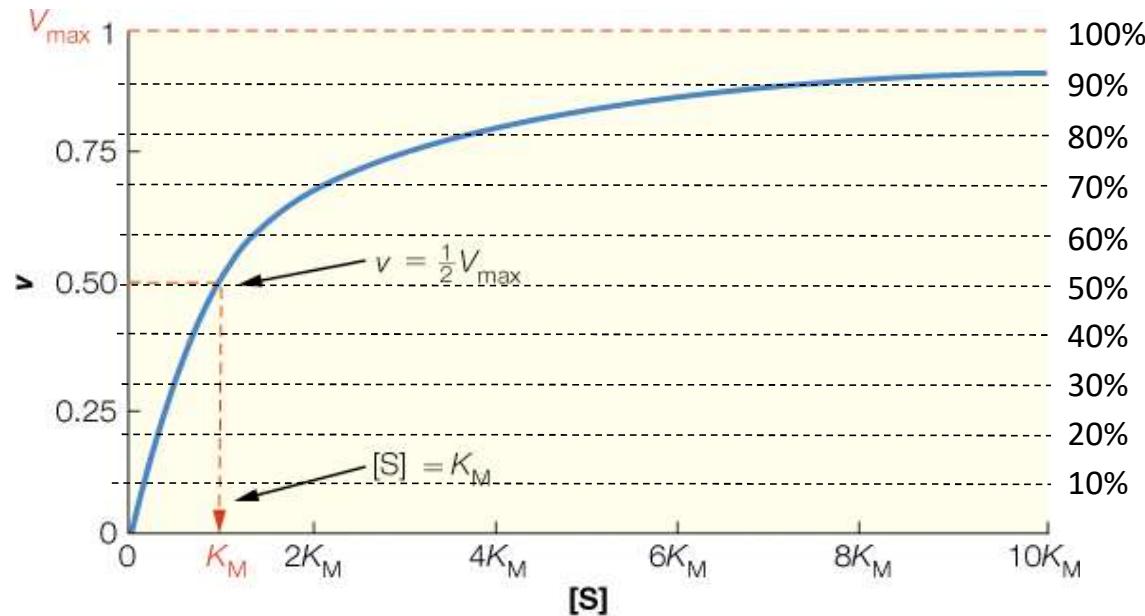
$$v = \frac{V_{\max} [S]}{[S] + K_m}$$

- V_{\max} is the maximum velocity at infinite substrate concentration.
- K_m is the Michaelis constant (units mol.L⁻¹)



Michaelis-Menten constants: K_m and the maximum rate, V_{max}

When $[S] = K_m$, $v = V_{max}/2$ or vice versa.



Kinetic data provides values for K_m and V_{max} : Double-Reciprocal (Lineweaver-Burk) Plot

- Rearranging the Michaelis-Menten relationship, we get a straight line equation:

$$\frac{1}{v} = \left(\frac{K_m}{V_{max}} \right) \frac{1}{[S]} + \frac{1}{V_{max}}$$

- This means that we can determine V_{max} and K_m graphically by plotting $(1/v)$ vs. $(1/[S])$
- Lineweaver-Burk double reciprocal plot (1934) (Prac 5)**

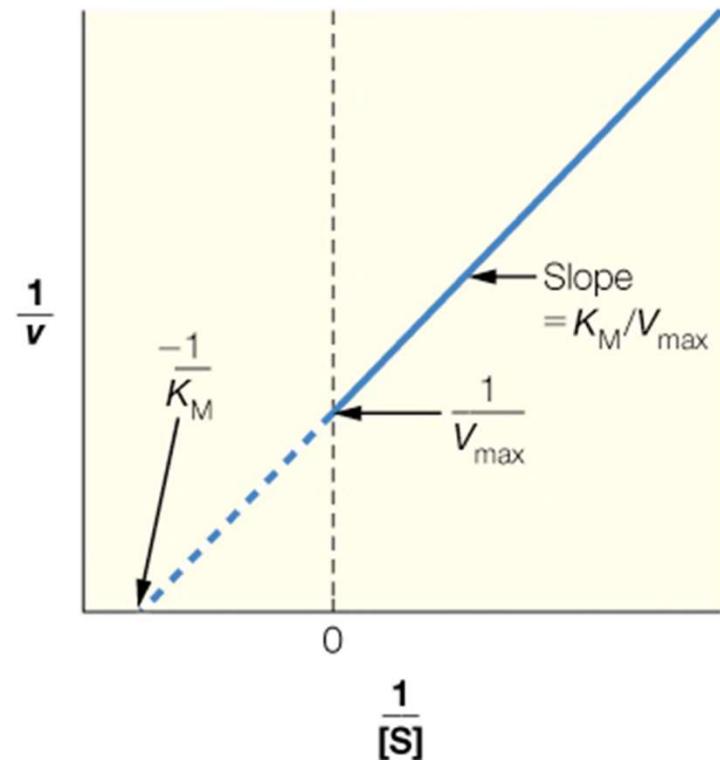


FIGURE 8.23 A Lineweaver–Burk plot.



The Michaelis constant, K_m

- $K_m = [S]$ when $v = V_{max}/2$
 - Units of mol.L⁻¹
- **Inversely** related to the **affinity** of a particular substrate for an enzyme
 - **Small K_m** = less substrate required for high reaction rate = **great affinity**
 - **High K_m** = more substrate required for high reaction rate = **low affinity**



The catalytic constant, k_{cat}

- enzyme property that measures its **internal speed**
- $k_{cat} = V_{max} / [E]$
 - **Turnover number** – the number of times a given site can turn substrate into product per second
 - Measures how **rapidly** the enzyme can operate
 - It has the unit s^{-1} (**Specific activity** is $k_{cat} / \text{mg of E}$)

The enzyme specificity constant

- k_{cat} / K_m
 - It has the units $\text{L.mol}^{-1}\text{s}^{-1}$
 - Measures enzyme efficiency
 - Determines which is the best substrate and how rapidly an enzyme will work at low [S]



Enzyme Kinetic Parameters

TABLE 8.5 Michaelis-Menten parameters for selected enzymes, arranged in order of increasing efficiency as measured by k_{cat}/K_M

| Enzyme | Reaction Catalyzed | $K_M(\text{mol/L})$ | $k_{\text{cat}}(\text{s}^{-1})$ | $k_{\text{cat}}/K_M [(\text{mol/L})^{-1}\text{s}^{-1}]$ |
|-------------------------|---|----------------------|---------------------------------|---|
| Chymotrypsin | Ac-Phe-Ala $\xrightarrow{\text{H}_2\text{O}}$ Ac-Phe + Ala | 1.5×10^{-2} | 0.14 | 9.3 |
| Pepsin | Phe-Gly $\xrightarrow{\text{H}_2\text{O}}$ Phe + Gly | 3×10^{-4} | 0.5 | 1.7×10^3 |
| Tyrosyl-tRNA synthetase | Tyrosine + tRNA \longrightarrow tyrosyl-tRNA | 9×10^{-4} | 7.6 | 8.4×10^3 |
| Ribonuclease | Cytidine 2', 3' cyclic phosphate $\xrightarrow{\text{H}_2\text{O}}$ cytidine 3'-phosphate | 7.9×10^{-3} | 7.9×10^2 | 1.0×10^5 |
| Carbonic anhydrase | $\text{HCO}_3^- + \text{H}^+ \longrightarrow \text{H}_2\text{O} + \text{CO}_2$ | 2.6×10^{-2} | 4×10^5 | 1.5×10^7 |
| Fumarase | Fumarate $\xrightarrow{\text{H}_2\text{O}}$ malate | 5×10^{-6} | 8×10^2 | 1.6×10^8 |

- K_m is unique for each enzyme-substrate pair.
- A low K_m value means the enzyme achieves maximal catalytic efficiency at low substrate concentrations.
- The larger the k_{cat} value, the greater the amount of product.



Enzyme Kinetics Summary

- ΔG determines if a reaction will occur:
 - $\Delta G = 0$, the reaction **is at equilibrium**
 - $\Delta G < 0$, the reaction **occurs spontaneously**
 - $\Delta G > 0$, the reaction **does not occur spontaneously.**
- When $\Delta G > 0$, this reaction can be powered by coupling it with a reaction that has $\Delta G < 0$ (such as $\text{ATP} \rightarrow \text{ADP} + \text{P}_i$)
- Michaelis constant, K_M for a particular enzyme reaction = [substrate] when $v = V_{\max}/2$.
- An enzyme's overall catalytic efficiency is expressed as k_{cat}/K_M .
- A Lineweaver–Burk plot can be used to experimentally plot kinetic data and calculate values for K_M and V_{\max} from the linear graph.



Enzymes 2: Inhibition and Control

Shoba Ranganathan

Applies Biosciences

T: 02 9850 6262; E: shoba.ranganathan@mq.edu.au

Objectives

- The active site
- Enzyme inhibition
 - At the active site
 - At another site
- Regulation of enzyme activity
 - Allosteric effectors
 - Chemical modifications

Textbook Chapter 8



MACQUARIE
University

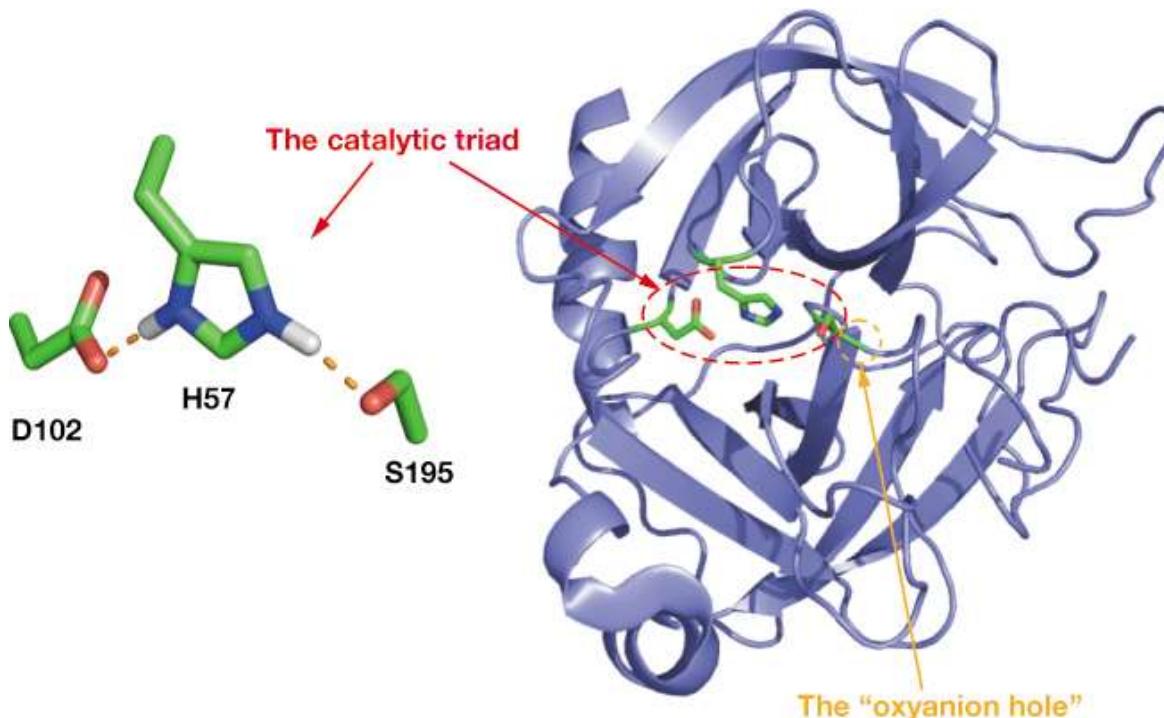
Module 2: Enzymes and Cell Biology

The active site

- The **active site** of a protein is where one or more **substrate** molecules bind, for the protein to carry out its function (defined in Chap. 7)
- Comprises a few residues critical for the protein or enzyme's function, usually located far apart in the amino acid sequence, but close in 3D space (e.g. serine protease catalytic triad).
- Mutating these residues either destroys or modifies enzyme function.
- Active site residues are localized to a specific substrate binding pocket by the folding of the enzyme 3D structure.
- Complements the transition state.



Serine proteases: digestive enzymes trypsin, chymotrypsin and elastase



- The three catalytic residues are absolutely essential for function
- H57, D102 and S195: far apart in the sequence
- But spatially localized and held in a specific geometry by the entire 3D structure.
- Histidine (H57) acts as a base under physiological conditions: i.e. neutral pH
- The OH of serine (S195) participates in the reaction
- Aspartate (D102) holds the histidine in place as well as hydrogen bonds with the peptide carbonyl of the substrate, after the recognition residue.

FIGURE 8.16 The structure of chymotrypsin and the serine protease catalytic triad.



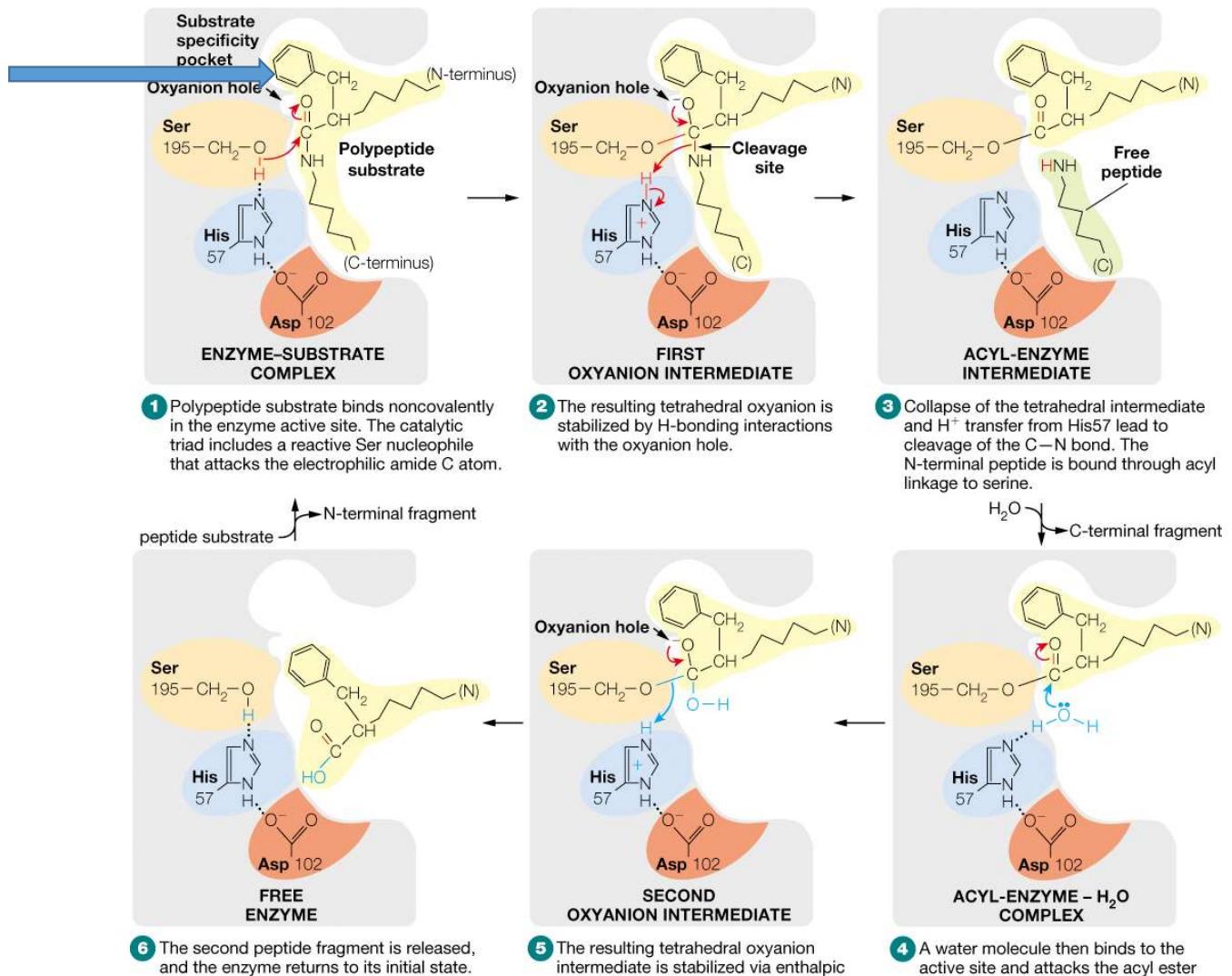


Figure 8.15 Catalysis of peptide bond hydrolysis by chymotrypsin

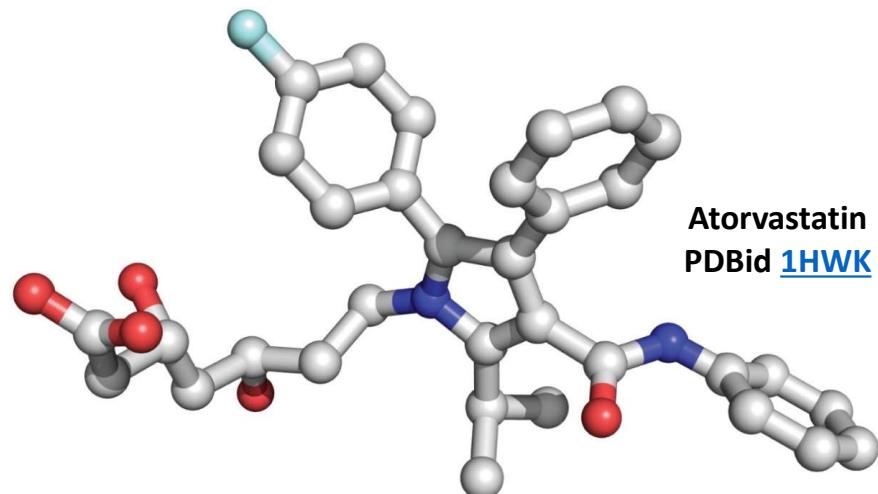
The enzyme active site - Summary

- The active site of an enzyme is a completely different environment to the surrounding solution
 - Can be more acidic or basic than the solution
 - Usually excludes solvent molecules and provides a hydrophobic environment
 - Restricts the orientation in which the substrate binds: i.e. stereospecificity
 - Possibility for electrophilic or nucleophilic functional groups
 - Reaction with the enzyme is possible but the enzyme is regenerated.
 - Critical for inhibitor design in biotech and drug design.



Enzyme inhibition: Atorvastatin (Lipitor®)

- One of the most commonly prescribed drugs for lowering cholesterol
- Binds to and **inhibits** the activity of the **enzyme** HMG-CoA reductase



Drugs, Toxins, and Enzymatic Activity

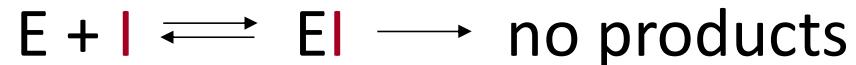
- Many prescription drugs are enzyme inhibitors
 - The development of some drugs has been advanced by kinetic studies, as well as structural studies of the target enzymes
 - Also, the effects of some natural and synthetic toxins are the result of enzyme inhibition or inactivation
 - There are two classes of enzyme inhibitors:
 - reversible inhibitors (noncovalently bound)
 - irreversible inhibitors (covalently bound)



Enzyme inhibition



- An enzyme (E) binds the substrate (S) reversibly forming the ES complex, which then forms product (P), with recovery of the enzyme.
- An inhibitor (I) disrupts the above reaction by a number of mechanisms, involving the inhibitor binding in some way (reversibly or irreversibly) to the enzyme (E)
 - The inhibitor can bind strongly to the **active site** and prevent the substrate (S) from binding, **or**
 - It can bind to **some other site** on the enzyme and affect the reaction indirectly.

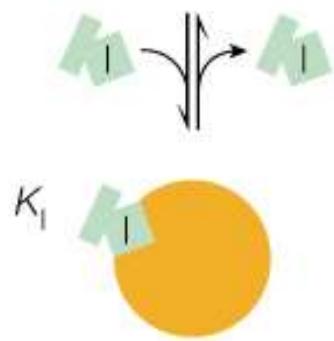
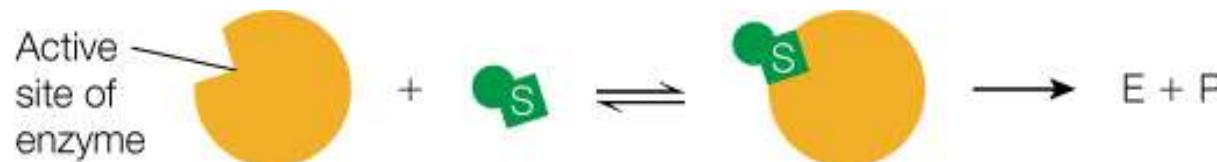


Reversible inhibition

- While in a reversible inhibition the inhibitor interacts with an enzyme noncovalently, in an irreversible reaction the inhibitor is bound covalently
- There are at least three different mode of reversible inhibition:
 - 1) competitive inhibition
 - 2) uncompetitive inhibition
 - 3) mixed or noncompetitive inhibition



Competitive inhibition



$$[E]_t = [E]_{\text{Free}} + [\text{ES}] + [\text{EI}]$$

Total enzyme Free enzyme Enzyme bound to substrate Enzyme bound to inhibitor

- Both substrate (S) and inhibitor (I) can fit the enzyme's active site.
- The substrate can be processed, but not the inhibitor



Competitive Inhibition

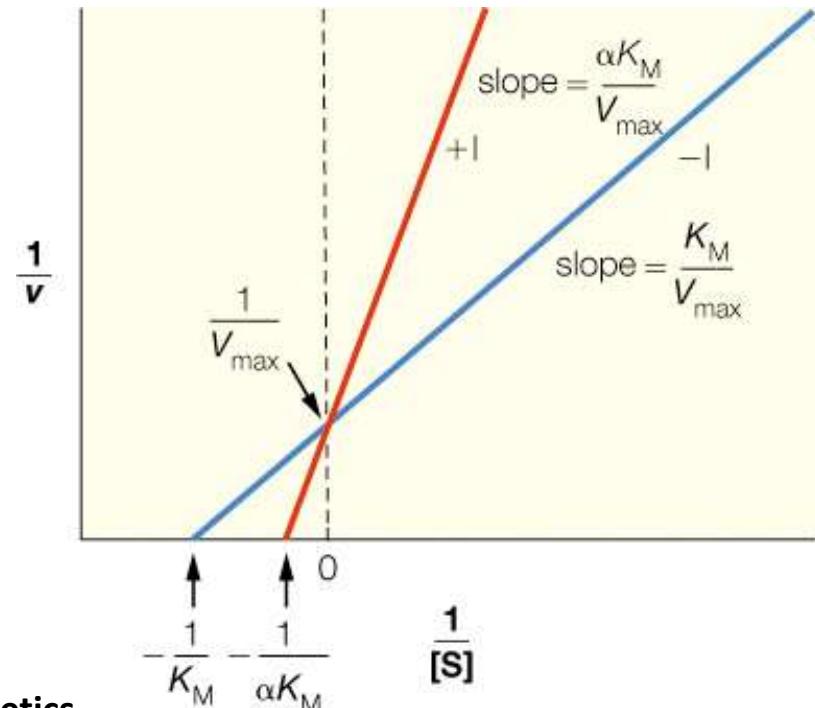
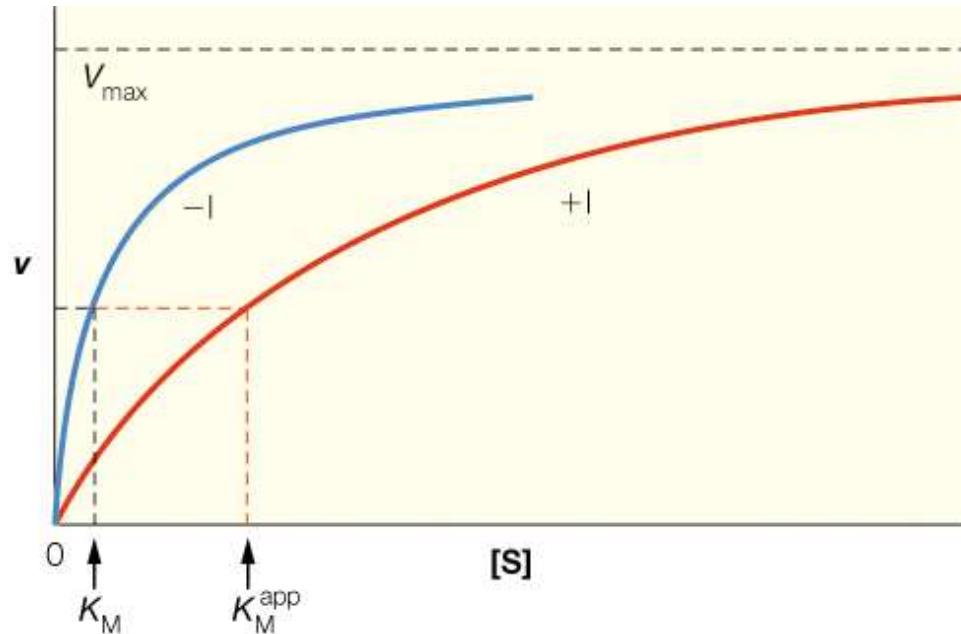


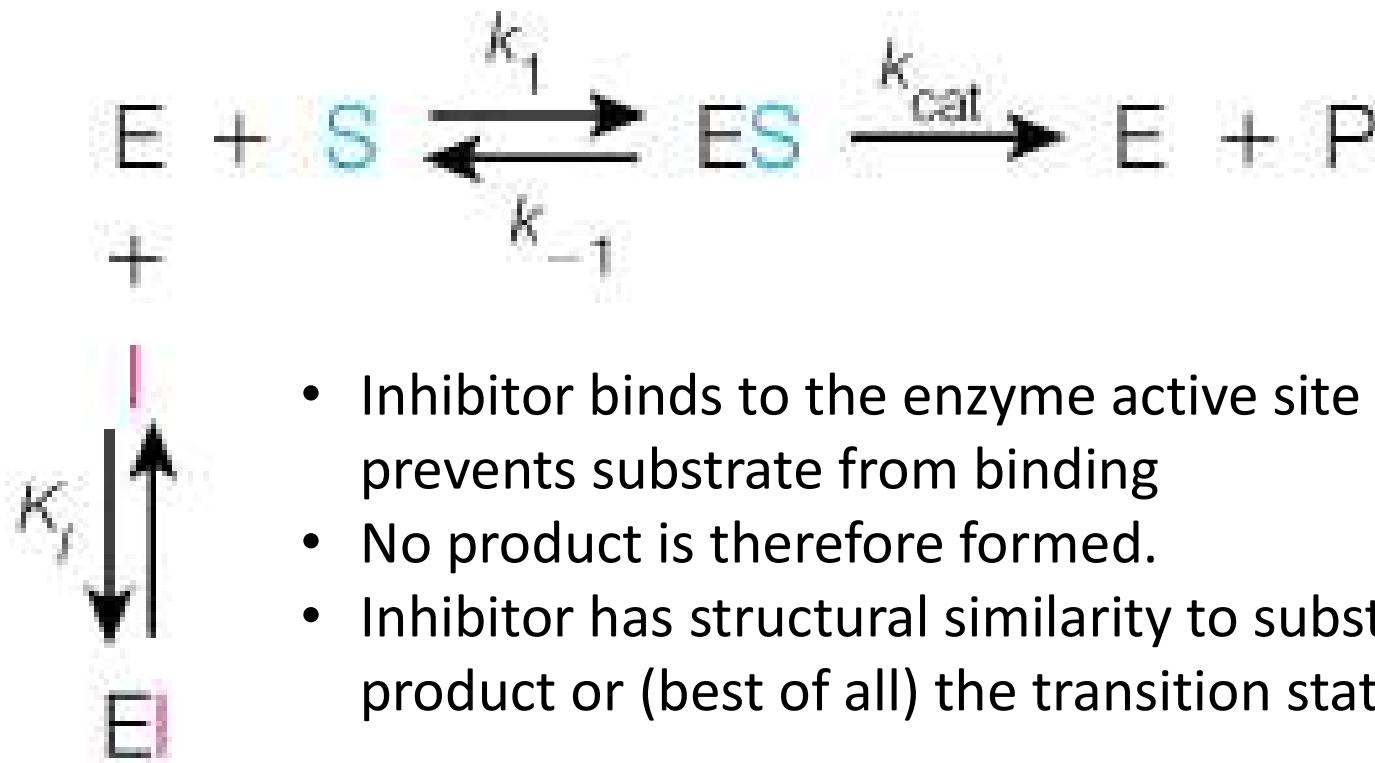
FIGURE 8.25 Effects of competitive inhibition on enzyme kinetics.

- Inhibitor binds to the **active site** and reduces amount of free enzyme available for catalysis by **competing with substrate**
 - K_m changes but V_{max} remains constant

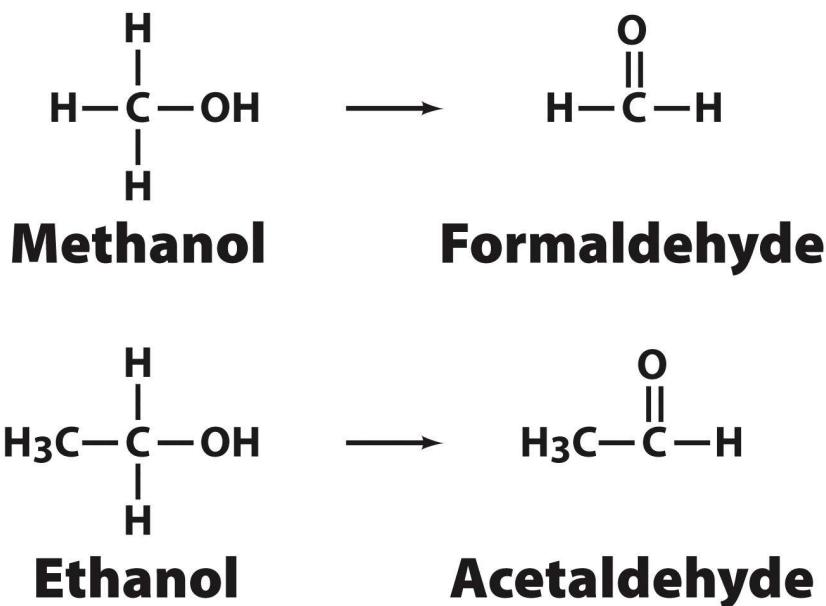
The apparent K_M is increased, while the V_{max} is unchanged



Competitive Enzyme Inhibition



Competitive Inhibition: Ethanol Treatment of Methanol Poisoning



- Alcohol dehydrogenase (ADH) is the enzyme.
- Ethanol competes for the active site with methanol.
- Methanol is harmlessly excreted *via* urine, instead of getting converted to toxic formaldehyde.



Malonate competitively inhibits succinate dehydrogenase (citric acid cycle)

| Substrate | Product | Competitive inhibitor |
|--|---|--|
| $\begin{array}{c} \text{COO}^- \\ \\ \text{CH}_2 \\ \\ \text{CH}_2 \\ \\ \text{COO}^- \end{array}$ | $\begin{array}{c} \text{COO}^- \\ \\ \text{CH} \\ \\ \text{HC} \\ \\ \text{COO}^- \end{array}$ | $\begin{array}{c} \text{COO}^- \\ \\ \text{CH}_2 \\ \\ \text{COO}^- \end{array}$ |

$\xrightarrow{\text{SDH}}$

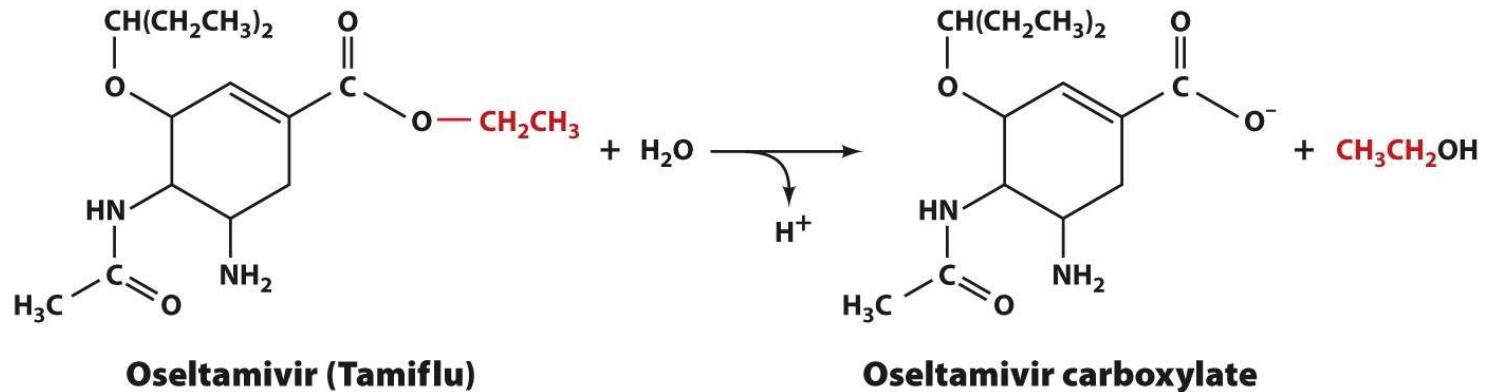
↓
2H

Succinate Fumarate Malonate

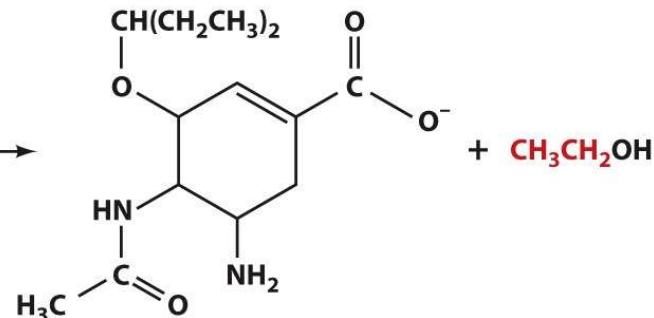
Note: structural similarity between S and I



Drugs are designed to fit the target enzyme's active site by competitive inhibition

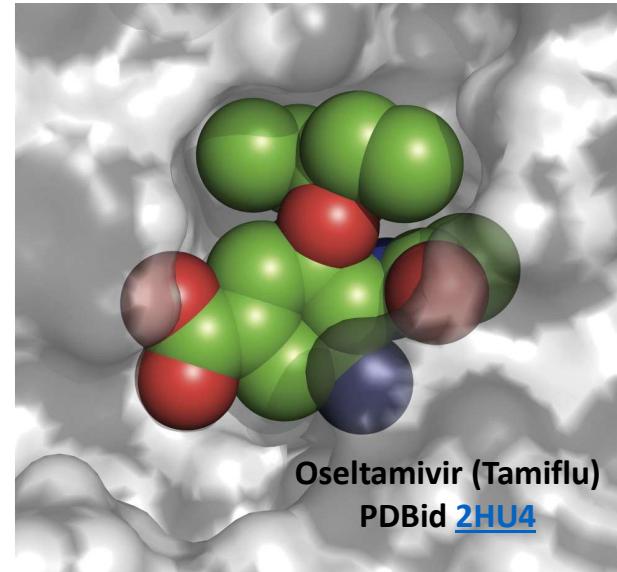


Oseltamivir (Tamiflu)

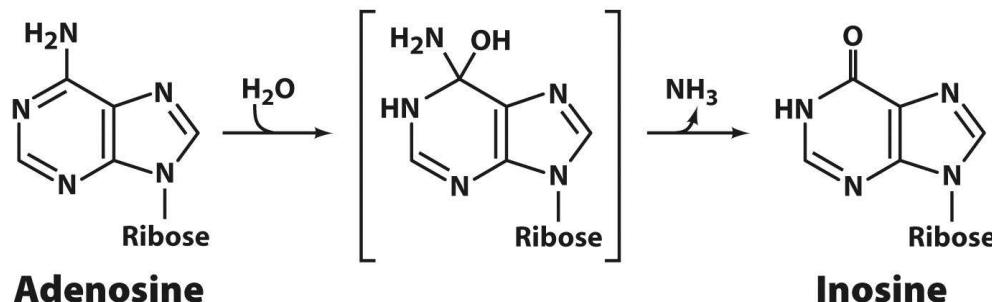


Oseltamivir carboxylate

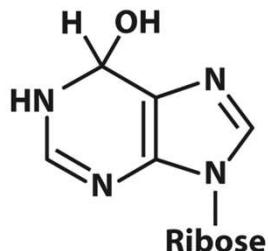
- Oseltamivir (Tamiflu®)-Avian Flu Neuraminidase Complex
- Neuraminidase catalyses the hydrolysis of neuraminic acid (sialic acid) to help viral particles escape from the host cell surface



Adenosine Deaminase: Transition State Analog Inhibitor for Leukemia Treatment



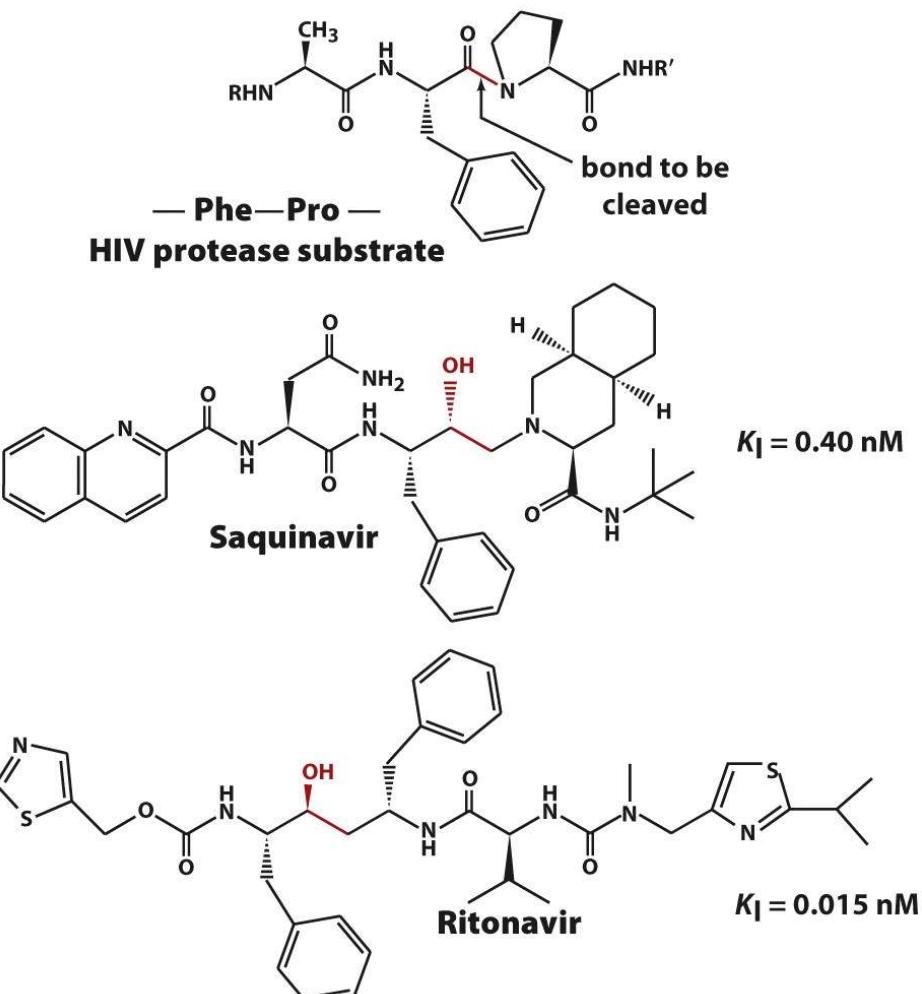
Unnumbered 12 p369
© 2013 John Wiley & Sons, Inc. All rights reserved.

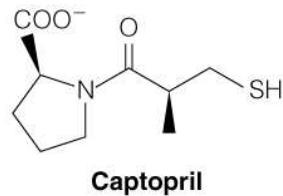
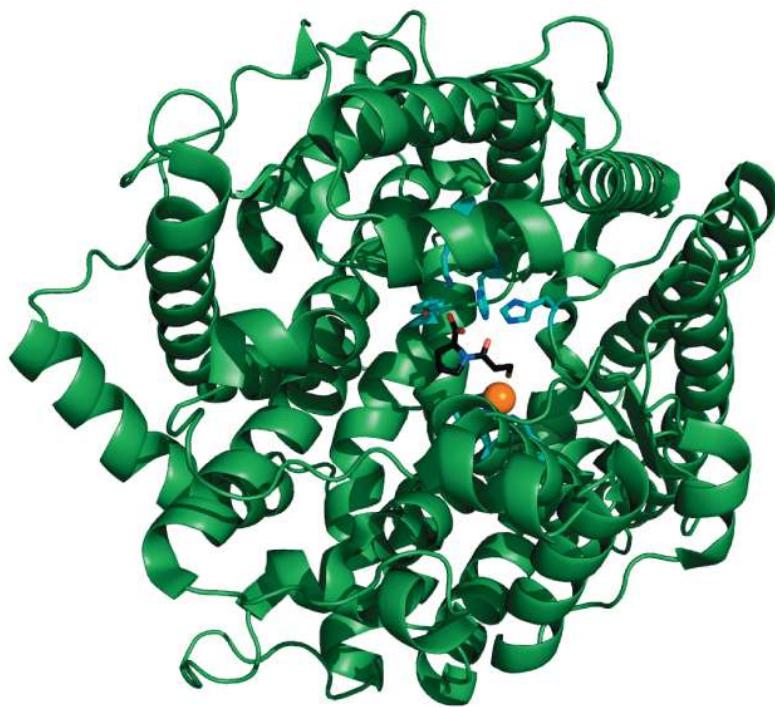


1,6-Dihydroinosine

Unnumbered 12 p370
© 2013 John Wiley & Sons, Inc. All rights reserved.

HIV Enzyme Inhibitors





$$K_I = 1.7 \text{ nM} \text{ for captopril}$$
$$K_M = 52 \mu\text{M} \text{ for angiotensin I}$$

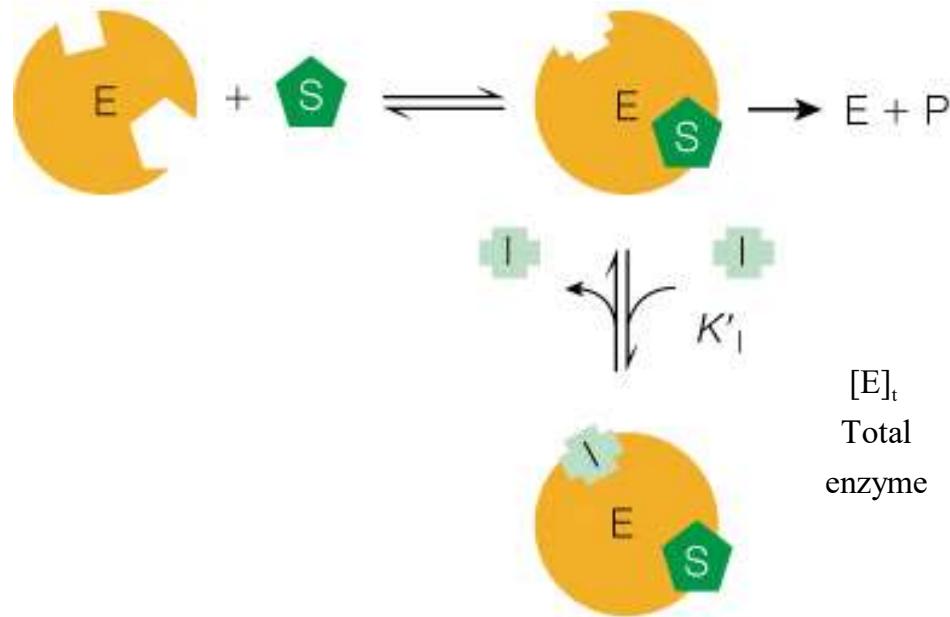
(a) A cartoon rendering of human ACE is shown in green with captopril (black) in sticks. The active site Zn^{2+} ion is shown as an orange sphere, and the side chains that bind captopril are shown as cyan sticks.



ACE inhibitor

- Angiotensin I is a peptide hormone that causes vasoconstriction and an increase in blood pressure.
Asp-Arg-Val-Tyr-Ile-His-Pro-Phe-His-Leu
- It is cut after Pro by the angiotensin converting enzyme, ACE to form angiotensin II.
- ACE inhibitors prevent congestive heart failure.

Figure 8.26a, Captopril is a competitive inhibitor of angiotensin-converting enzyme.



Uncompetitive inhibition

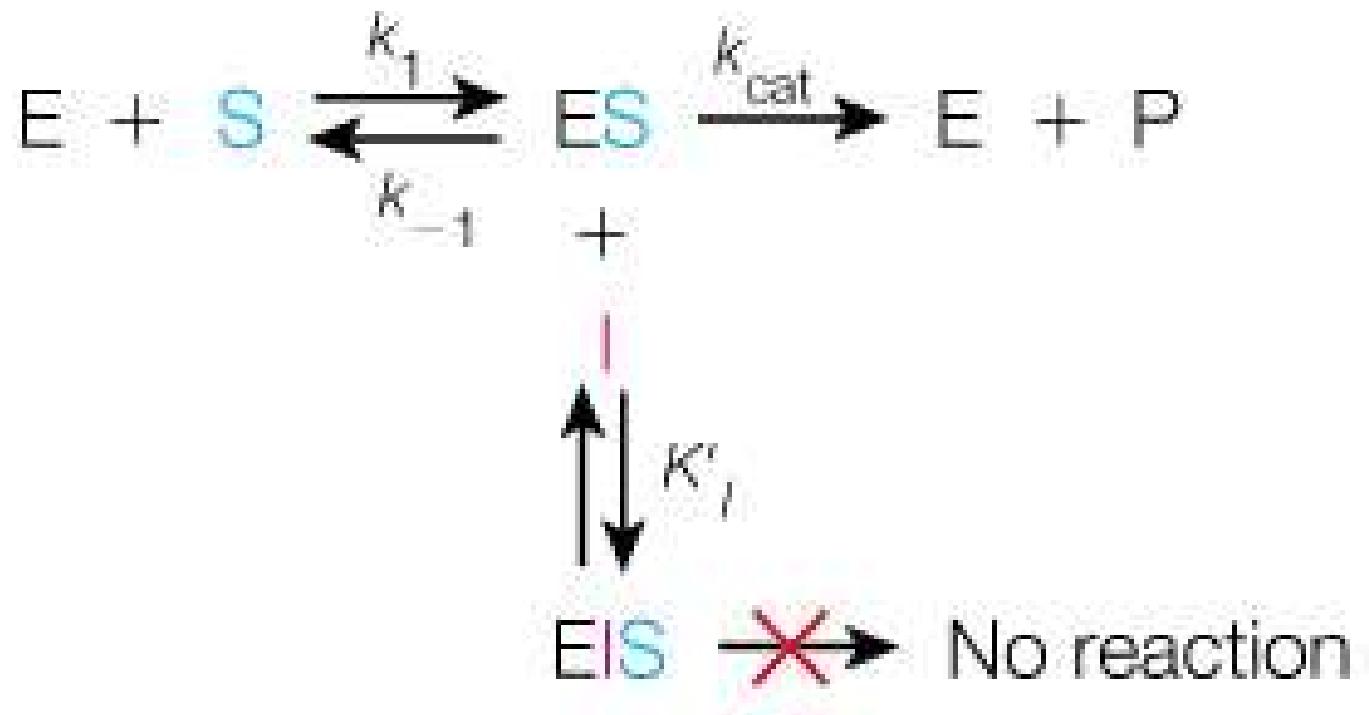
$$[E]_t = [E]_{\text{Free}} + [ES]_{\text{to substrate}} + [EI]_{\text{to inhibitor}}$$

Total enzyme Free enzyme Enzyme bound to substrate Enzyme bound to inhibitor

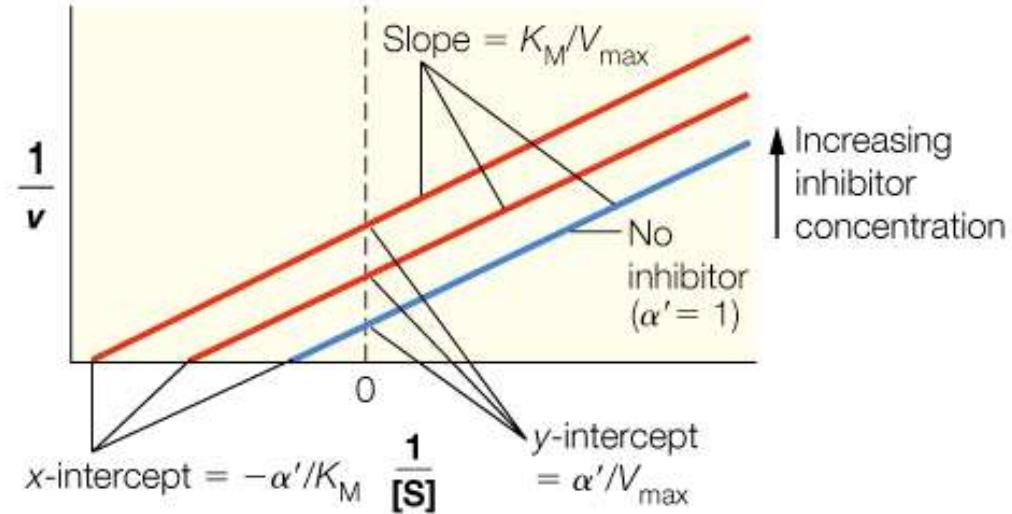
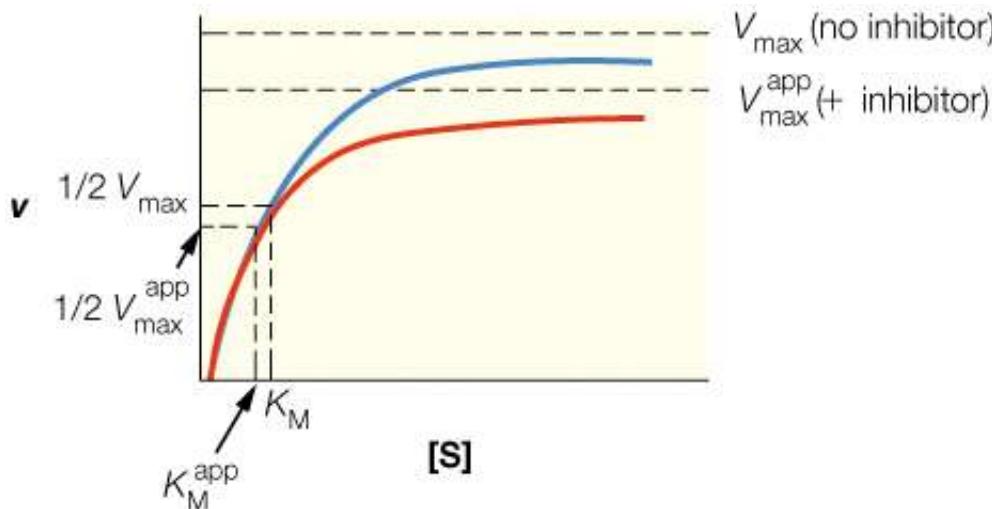
- The inhibitor binds at a site on the enzyme different from that of the substrate, thereby diminishing the enzyme's catalytic activity
- This is a **regulatory site** where a bound **effector** acts by an **allosteric** mechanism.
- Inhibitor I binds to enzyme after ES has formed
 - I prevents product formation
 - Dead-end complex



Uncompetitive Enzyme Inhibition



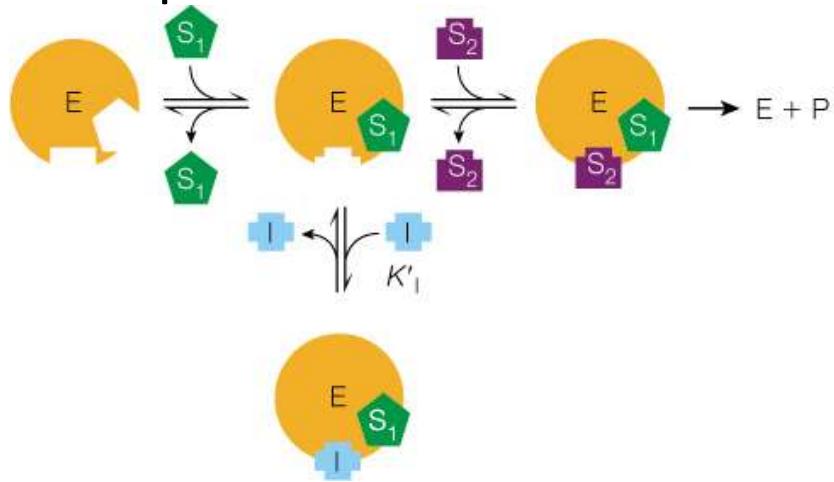
Uncompetitive Inhibition



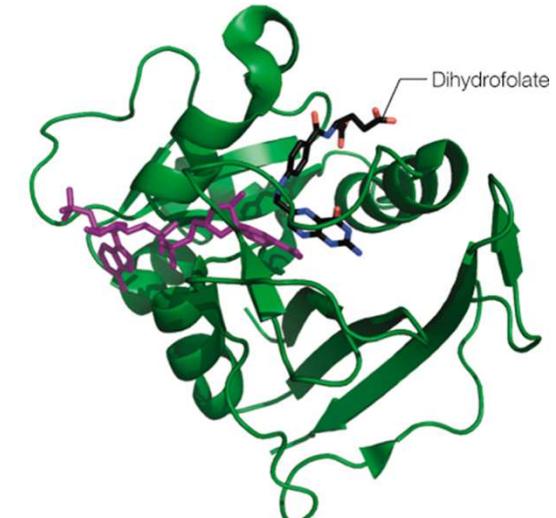
- Both V_{\max} and K_m change in this case.
- Slope (K_m/V_{\max}) is the same
- Both, the apparent K_M and apparent V_{\max} , are decreased



Uncompetitive Inhibition Example



(a) Methotrexate (black sticks) binds tightly to the complex of dihydrofolate reductase (green cartoon) and NADPH (magenta sticks). NADPH is a phosphorylated derivative of NADH. PDB ID: 1u72.

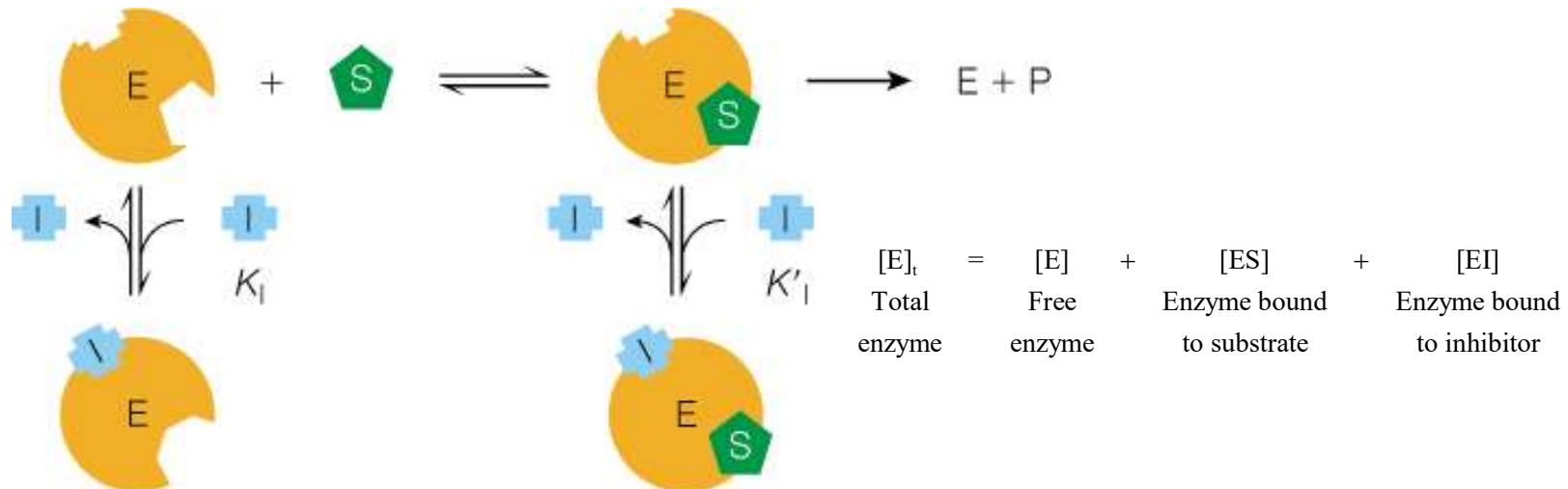


(b) Dihydrofolate reductase bound to both its normal substrates NADPH (magenta sticks) and dihydrofolate (black sticks). PDB ID: 2w3m.

- Bisubstrate reaction with ordered substrate binding involving an inhibitor that is uncompetitive with respect to substrate S_1 and competitive with respect to substrate S_2
- Inhibition of dihydrofolate reductase by the anticancer drug methotrexate; uncompetitive for NADPH (S_1 , cofactor) binding and competitive for dihydrofolate (S_2) binding.



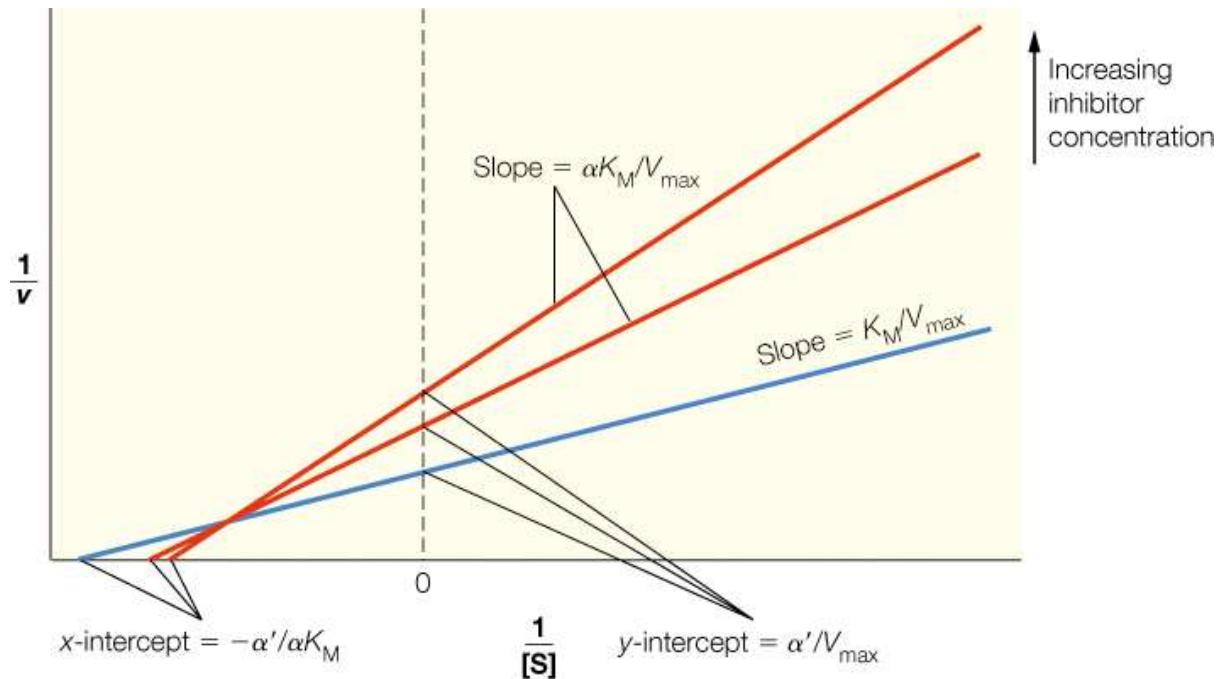
Mixed or noncompetitive inhibition



- The inhibitor binds at a site on the enzyme **other than the active site**, thereby diminishing the enzyme's catalytic activity
- Inhibitor binding occurs at any stage during catalysis
 - I does not compete with S for the active site
 - Both E and ES can bind I



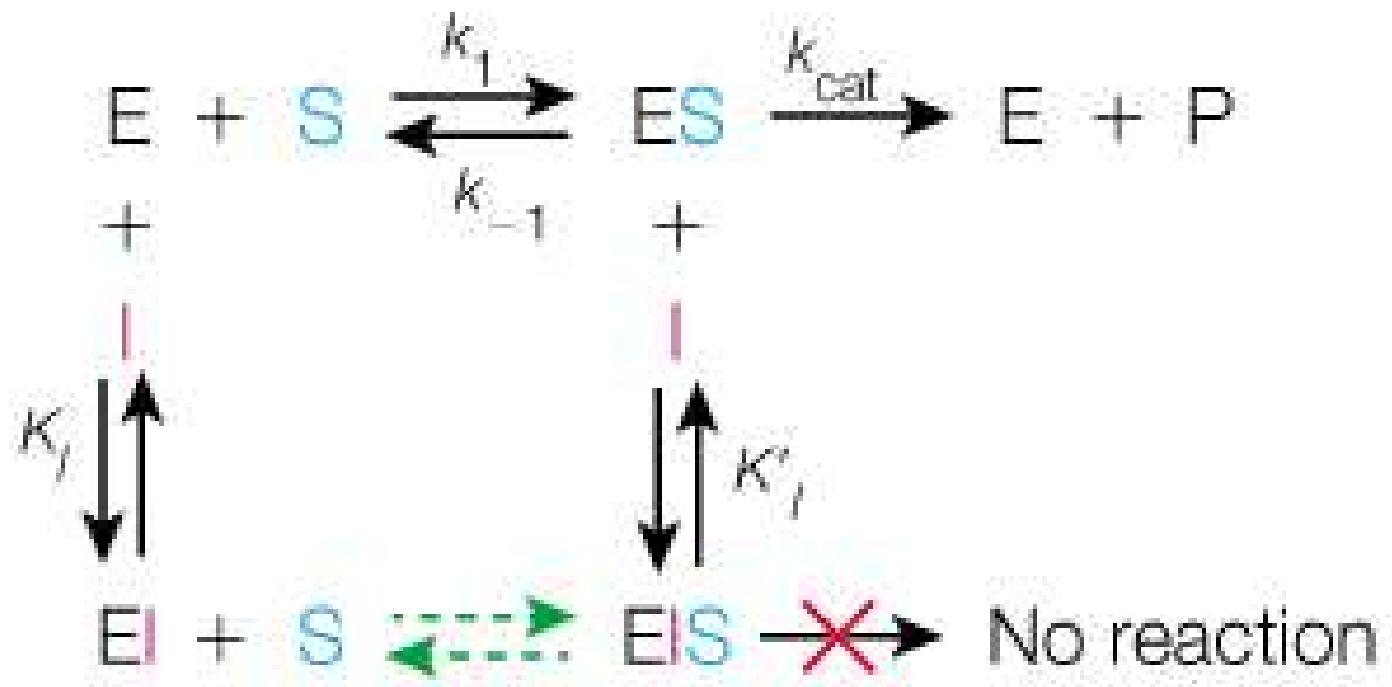
Mixed or Noncompetitive Inhibition



The apparent K_M is almost unchanged, while the apparent V_{max} is decreased



Mixed (Noncompetitive) Enzyme Inhibition



Enzyme Inhibitor Effects

| Type of Inhibition | Effect of Inhibitor |
|------------------------|---|
| None | None |
| Competitive | Increases K_M^{app} |
| Uncompetitive | Decreases K_M^{app} and V_{\max}^{app} |
| Mixed (noncompetitive) | Decreases V_{\max}^{app} ; may increase or decrease K_M^{app} |



Irreversible Inhibition

Covalent bonding to an enzyme's active site

- Example: diisopropyl fluorophosphate (DFP) binds to the active site serine of acetylcholinesterase, an enzyme involved in nerve conduction.
- Inhibition of this enzyme causes rapid paralysis

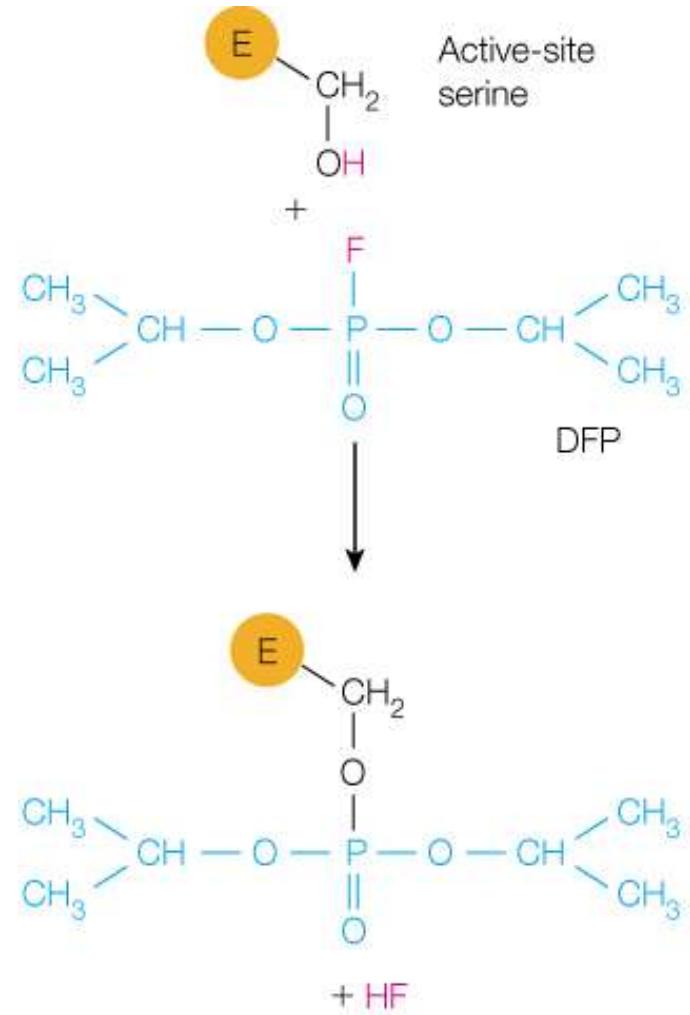


FIGURE 8.32 Irreversible inhibition by adduct formation.



Enzyme Inhibition Summary

- Enzyme inhibitors interact reversibly or irreversibly with an enzyme to alter its K_M and/or V_{max} values.
- A competitive inhibitor binds to the enzyme's active site and increases the apparent K_M for the reaction.
- An uncompetitive enzyme inhibitor affects catalytic activity such that both the apparent K_M and the apparent V_{max} decrease (but ratio remains the same).
- A mixed enzyme inhibitor alters both catalytic activity and substrate binding such that the apparent V_{max} decreases and the apparent K_M may increase or decrease slightly.
- Irreversible inhibition is by covalent bond formation.



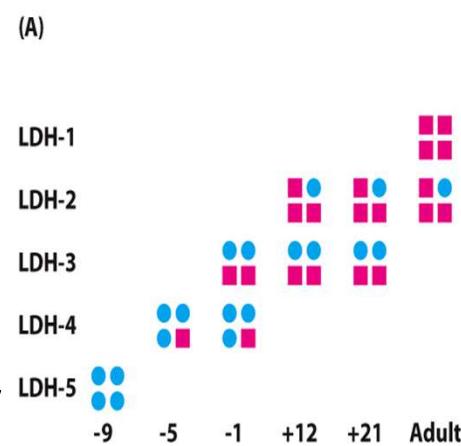
Enzyme activity can be controlled

- By enzyme availability: how much is present or produced?
Controlled by the cell and can change dramatically.
 - “Gene level control” can also lead to the generation of multiple isomeric forms called isozymes
- By allosteric mechanisms:
 - Structural changes affecting substrate binding
- By covalent modification
 - Usually phosphorylation (by kinases) and dephosphorylation (by phosphatases) of specific residues: Ser, Thr or Tyr.
 - Can activate or deactivate enzymes: **on or off switches**



Control by Multiple Forms of Enzymes

- Isoenzymes or isozymes are enzymes that are encoded by different genes.
- They catalyze the same reaction but display different regulatory properties.
- They may be expressed in a tissue-specific or developmentally specific pattern.
- The appearance of certain isozymes in the blood is a sign of tissue damage.



(B)

| | Heart | Kidney | Red blood cell | Brain | Leukocyte | Muscle | Liver |
|-------------------------------|-------|--------|-------------------|-------|-----------|--------|-------|
| H ₄ | — | — | — | — | — | — | — |
| H ₃ M | — | — | — | — | — | — | — |
| H ₂ M ₂ | — | — | — | — | — | — | — |
| HM ₃ | — | — | — | — | — | — | — |
| M ₄ | — | — | — | — | — | — | — |

Figure 10.15
Biochemistry, Eighth Edition
© 2015 Macmillan Education

Control by Multiple Forms of Enzymes

Studies of individual enzyme molecules suggest that some enzymes (isozymes) may exist in multiple conformations that are in equilibrium.

These different conformations may have different catalytic or regulatory properties – more in amino acid metabolism.

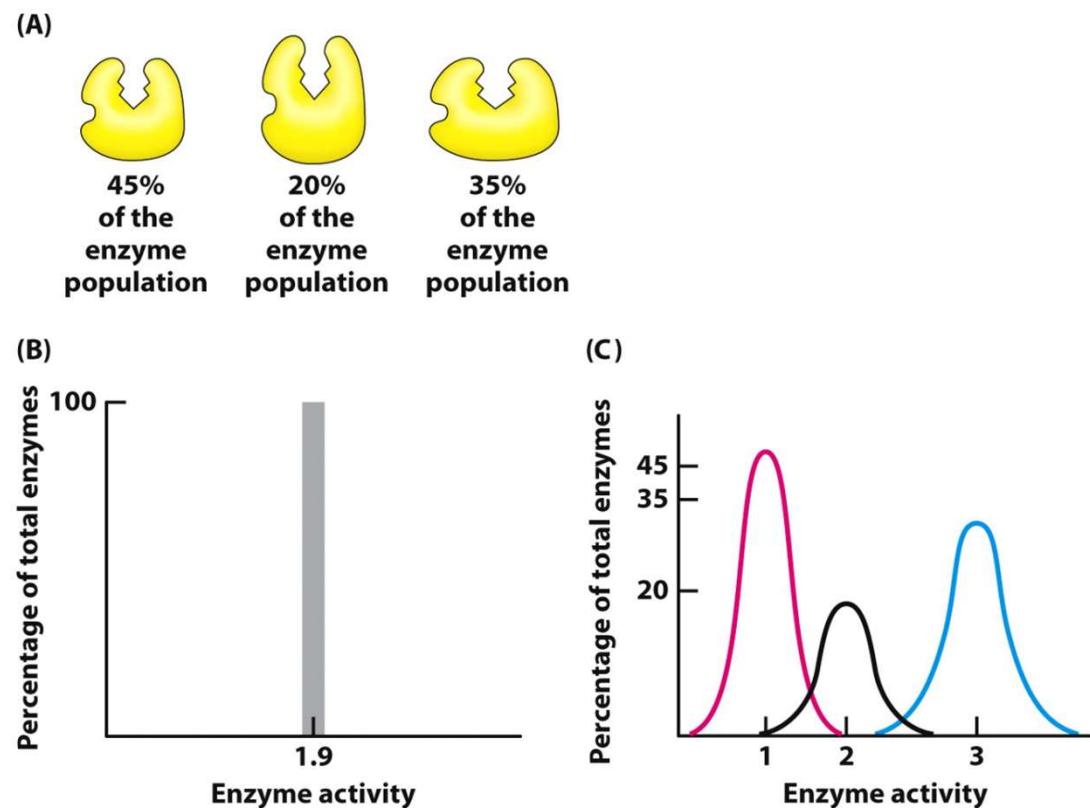
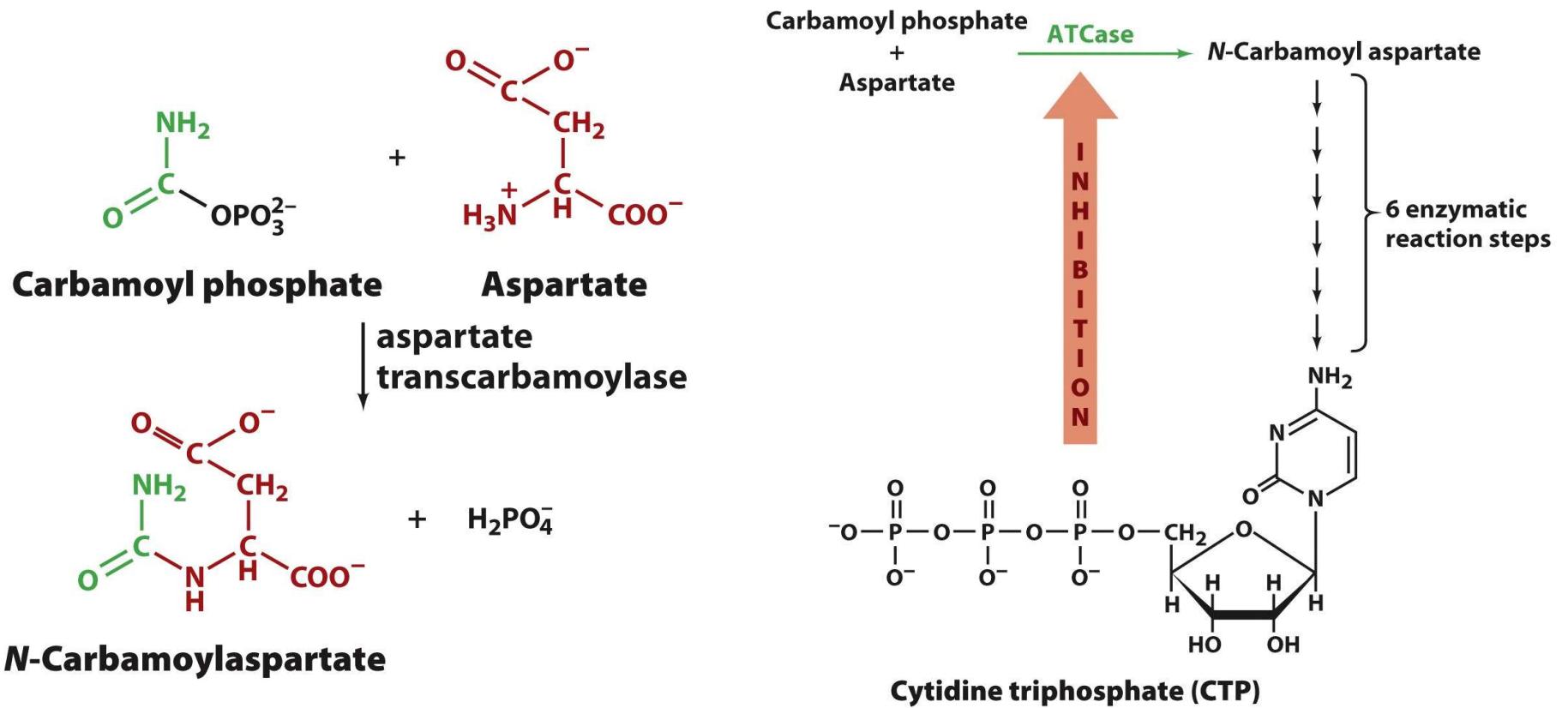


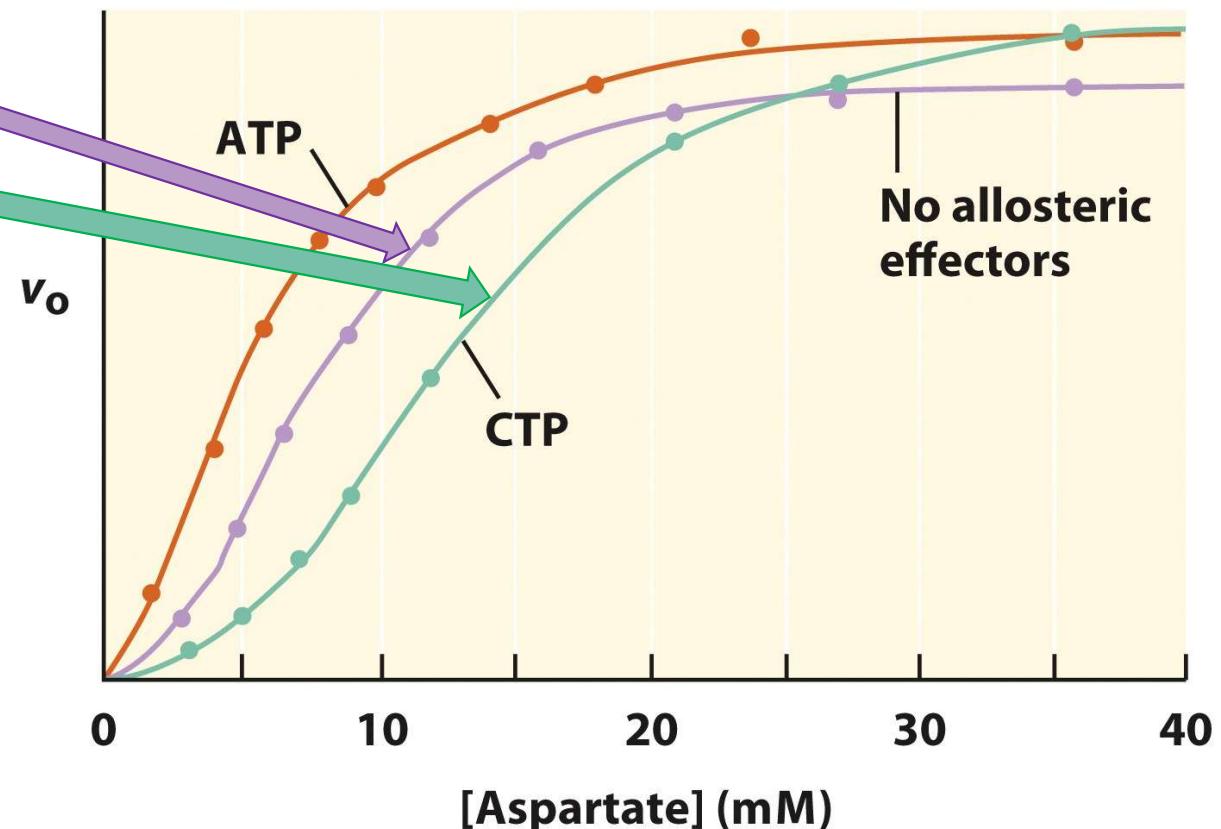
Figure 8.34
Biochemistry, Seventh Edition
© 2012 W. H. Freeman and Company

Allosteric control: Pyrimidine Biosynthesis: Aspartate transcarbamoylase (ATCase) Feedback Inhibition



Allosteric Effectors: ATCase Reaction

- Reaction velocity curve changes from hyperbolic to sigmoidal
- ATP is a positive effector, while CTP is a negative effector
- Final product in a multi-step reaction inhibits the enzyme catalysing the first step
- Conserves the cell's resources
- Effector binds to the allosteric site, not active site
- *Another example in Prac 5*



Control by Reversible Covalent Modification

TABLE 10.1 Common covalent modifications of protein activity

| Modification | Donor molecule | Example of modified protein | Protein function |
|------------------|---------------------------------------|-----------------------------|--|
| Phosphorylation | ATP | Glycogen phosphorylase | Glucose homeostasis; energy transduction |
| Acetylation | Acetyl CoA | Histones | DNA packing; transcription |
| Myristoylation | Myristoyl CoA | Src | Signal transduction |
| ADP ribosylation | NAD ⁺ | RNA polymerase | Transcription |
| Farnesylation | Farnesyl pyrophosphate | Ras | Signal transduction |
| γ-Carboxylation | HCO ₃ ⁻ | Thrombin | Blood clotting |
| Sulfation | 3'-Phosphoadenosine-5'-phosphosulfate | Fibrinogen | Blood-clot formation |
| Ubiquitination | Ubiquitin | Cyclin | Control of cell cycle |

Table 10.1
Biochemistry, Eighth Edition
© 2015 Macmillan Education



MACQUARIE
University

Module 2: Enzymes and Cell Biology

Control by Covalent Modification: Phosphorylation

Protein kinases modify proteins by attaching a phosphate to a serine, threonine, or tyrosine residue. ATP serves as the phosphate donor.

TABLE 10.2 Examples of serine and threonine kinases and their activating signals

| Signal | Enzyme |
|---|--|
| Cyclic nucleotides | Cyclic AMP-dependent protein kinase Cyclic GMP-dependent protein kinase |
| Ca ²⁺ and calmodulin | Ca ²⁺ -calmodulin protein kinase Phosphorylase kinase or glycogen synthase kinase 2 |
| AMP | AMP-activated kinase |
| Diacylglycerol | Protein kinase C |
| Metabolic intermediates and other "local" effectors | Many target-specific enzymes, such as pyruvate dehydrogenase kinase and branched-chain ketoacid dehydrogenase kinase |

Source: Information from D. Fell, *Understanding the Control of Metabolism* (Portland Press, 1997), Table 7.2.

Table 10.2
Biochemistry, Eighth Edition
© 2015 Macmillan Education



MACQUARIE
University

Module 2: Enzymes and Cell Biology

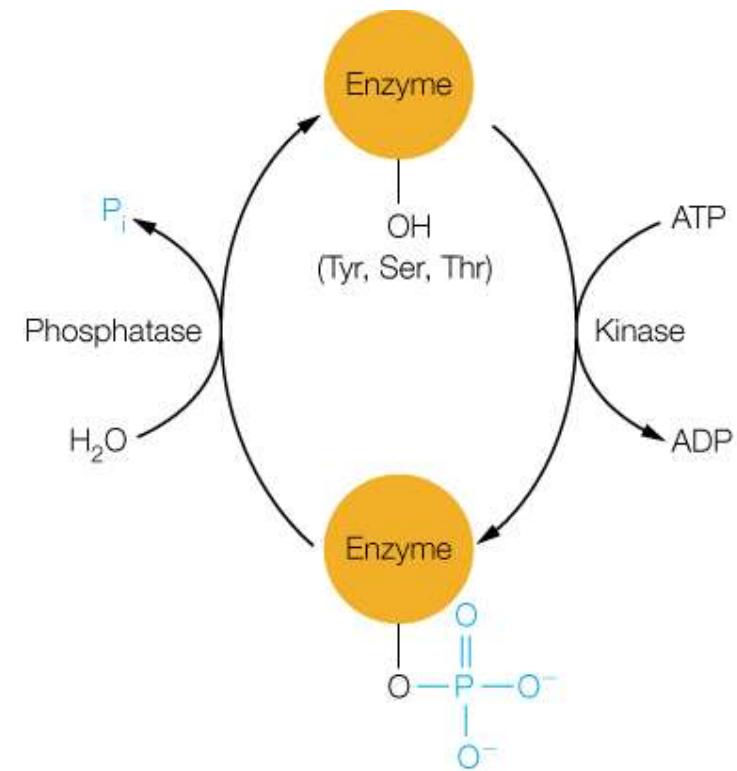


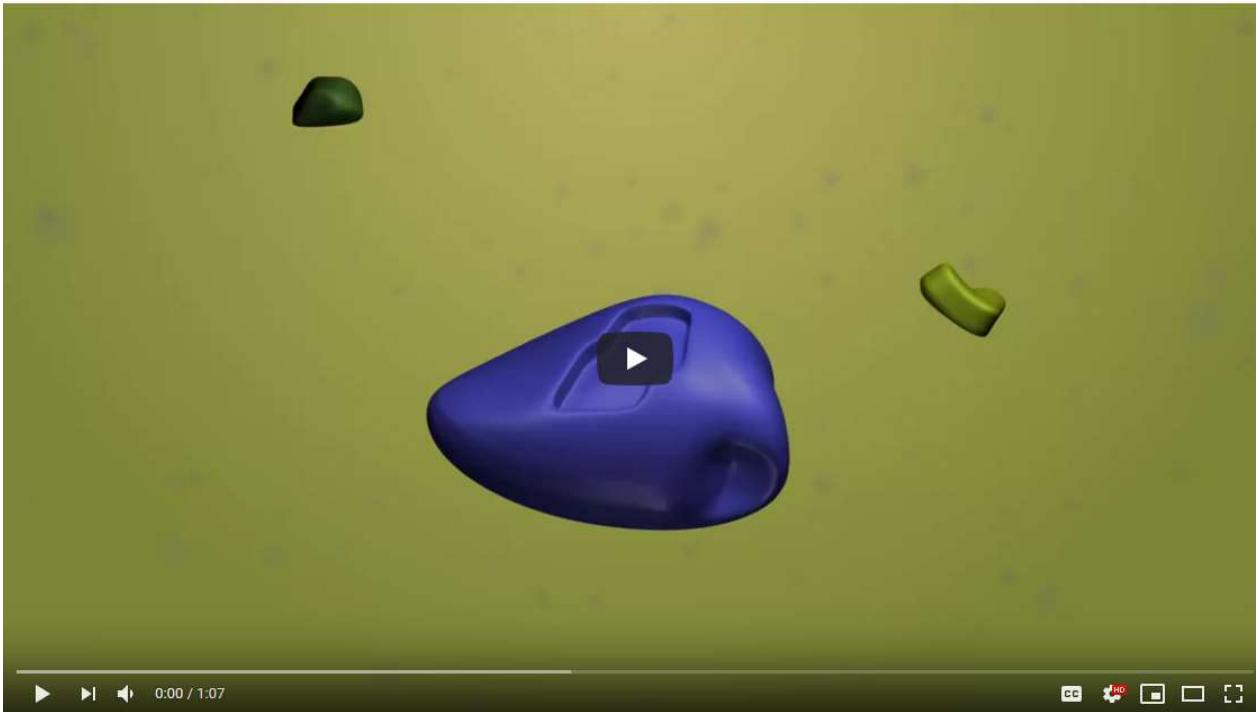
FIGURE 8.41 Reversible covalent modification by kinases/phosphatases.

Control of Enzyme Activity Summary

- Activities of enzymes can be regulated in various ways by: e.g., gene level control, allosteric mechanisms, or reversible/irreversible covalent protein modification
- Allosteric effectors bind to multisubunit enzymes such as aspartate transcarbamoylase (ATCase), thereby inducing cooperative conformational changes that alter the enzyme's catalytic activity.
- An example of reversible modification is the phosphorylation and dephosphorylation of enzymes by kinases and phosphatases, respectively
 - Phosphorylation and dephosphorylation of an enzyme such as glycogen phosphorylase can control its activity by shifting the equilibrium between more active and less active conformations.



Enzyme function and inhibition



<https://www.youtube.com/watch?v=PILzvT3spCQ>



MACQUARIE
University

Module 2: Enzymes and Cell Biology

Enzyme mechanisms – how enzymes work

- Lock and key model (Fischer, 1894)
 - The substrate fits into the active site of the enzyme like a key fits into a lock (historic).
- **Induced fit model**
 - The active site *does not quite fit* the substrate.
 - So the active site is changed to accommodate the substrate.
 - ✓ This flexibility can help the substrate enter the transition state and the product to leave.
 - ✓ Catalytically active antibodies are made against transition state analogues.
 - Instead of the enzyme, the substrate can be fitted to the active site.

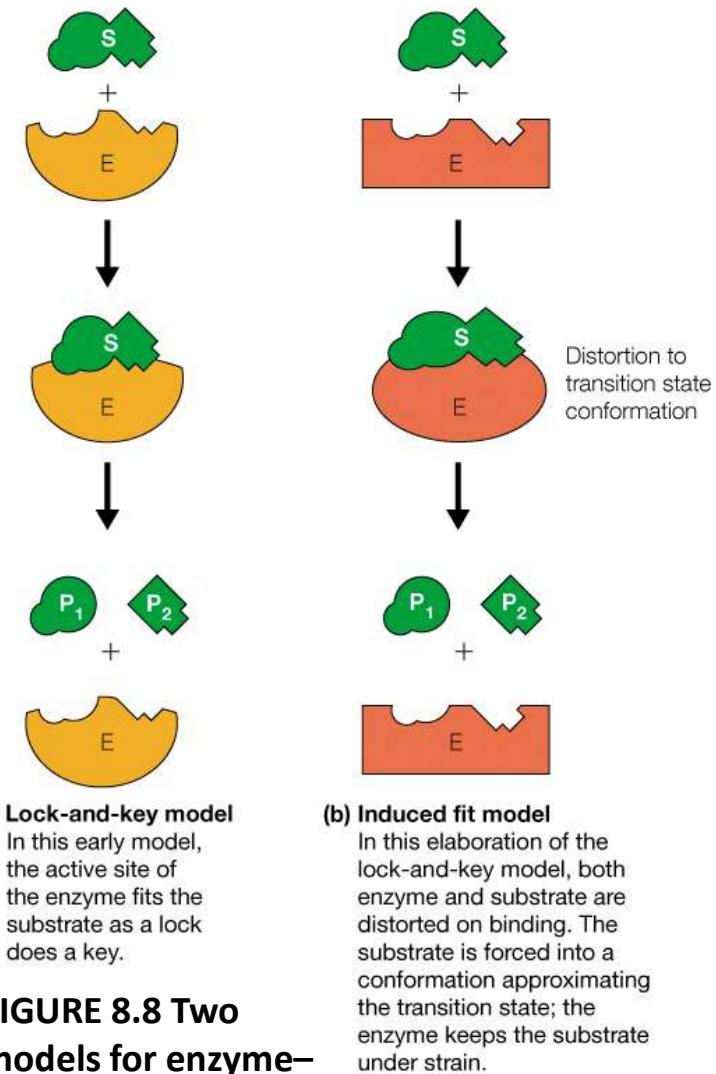


FIGURE 8.8 Two models for enzyme–substrate interaction.

Induced fit model from experiment

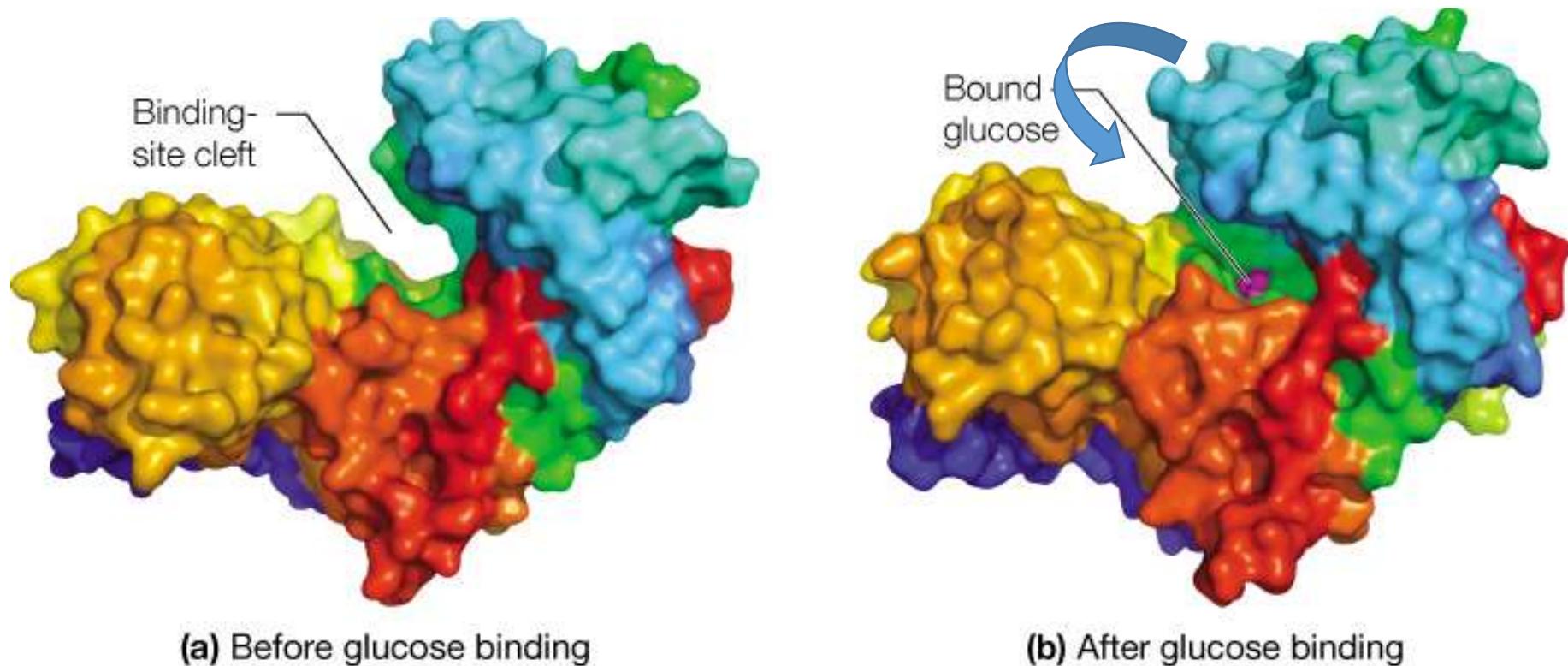


FIGURE 8.9 The induced conformational change in hexokinase.

