Refresh

Database
Characteristics
Relational database
Database collections
DNA sequence databases
Literature databases

Protein structure and function

Proteins

* Extremely versatile in their function

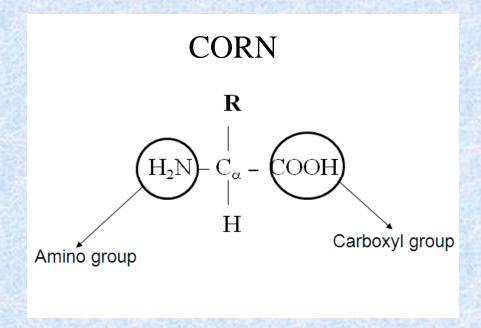
Nucleic acids

Carry genetic information

Carbohydrates

- Staple of human diet
- Structural and protective elements

Building blocks



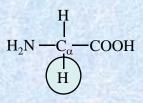
3 and 1 letter codes

Ala	Α	Methionine	Met	M
Cys	С	Asparagine	Asn	N
Asp	D	Proline	Pro	Р
Glu	E	Glutamine	Gln	Q
Phe	F	Arginine	Arg	R
Gly	G	Serine	Ser	S
His	Н	Threonine	Thr	T
lle				
Lys	K	Valine	Val	V
Leu	L	Tryptophan	Trp	W
		Tyrosine	Tyr	Υ
	Cys Asp Glu Phe Gly His Ile Lys	Cys C Asp D Glu E Phe F Gly G His H Ile I Lys K	Cys C Asparagine Asp D Proline Glu E Glutamine Phe F Arginine Gly G Serine His H Threonine Ile I Lys K Leu L Tryptophan	Cys C Asparagine Asn Asp D Proline Pro Glu E Glutamine Gln Phe F Arginine Arg Gly G Serine Ser His H Threonine Thr Ile I Lys K Leu L Tryptophan Trp

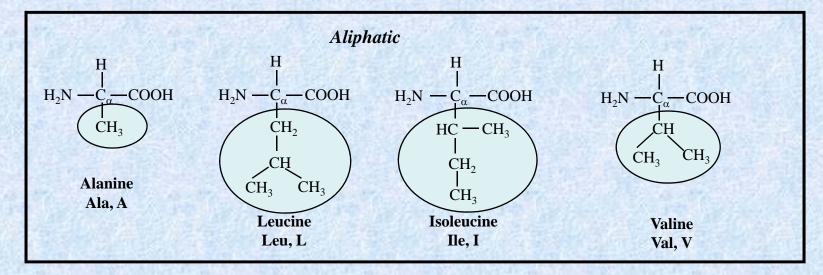
B, J, O, U, X, Z

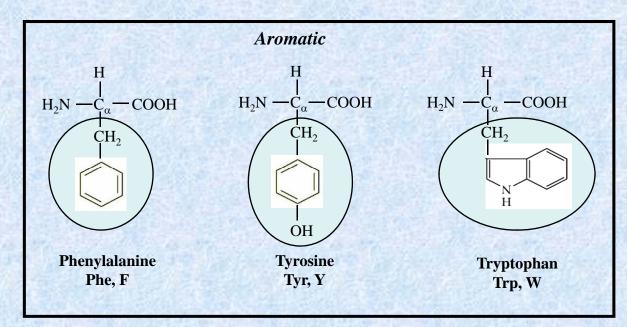
Hydrophobic residues

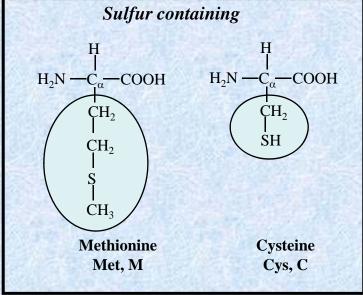
Hydrogen



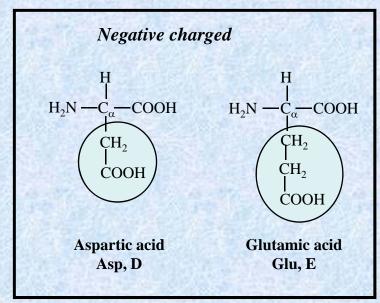
Glycine Gly, G

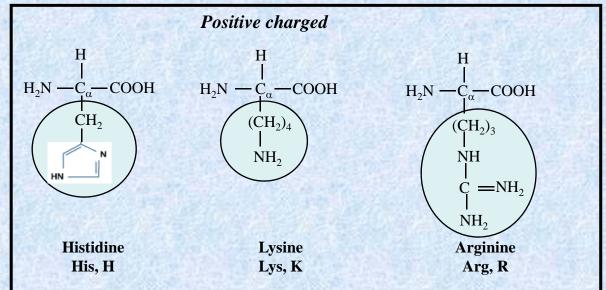


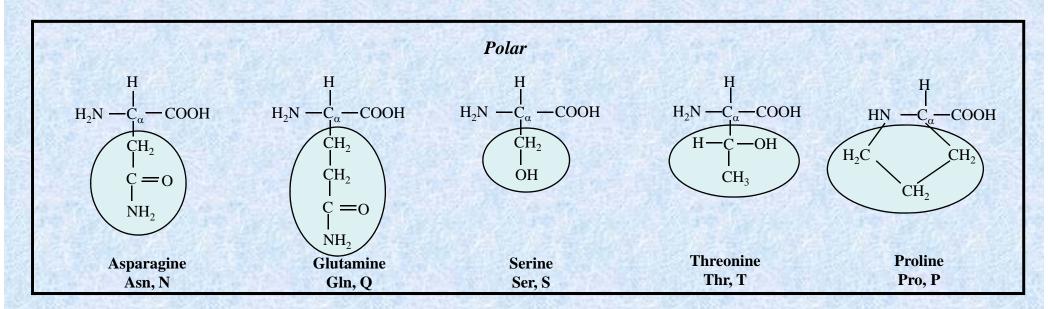




Hydrophilic residues





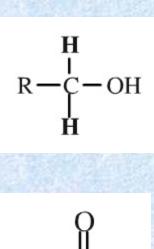


Notations

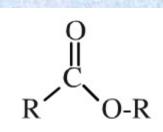
Greek alphabet.

```
    α alpha
    β beta
    γ gamma
    δ delta
    ε epsilon
    ζ zeta
    η eta
    θ theta
    φ phi
    χ chi
    ψ psi
    ω omega
```

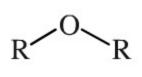
Chemical groups



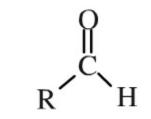
Alcohol



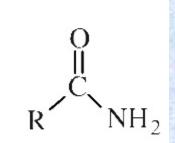
Ester



Ether



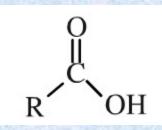
Aldehyde



Amide

R

Alkane



Acid

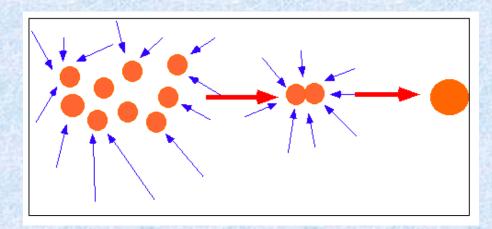
 $R-NH_2$

Amine

Aliphatic Aromatic

Amino acids: properties

1. Nonpolar (Hydrophobic; water hating; hydrocarbon alkyl groups (alkane branches) or aromatic (benzene rings)



2. Polar (hydrophilic; water loving): functional groups such as acids, amides, alcohols, and amines

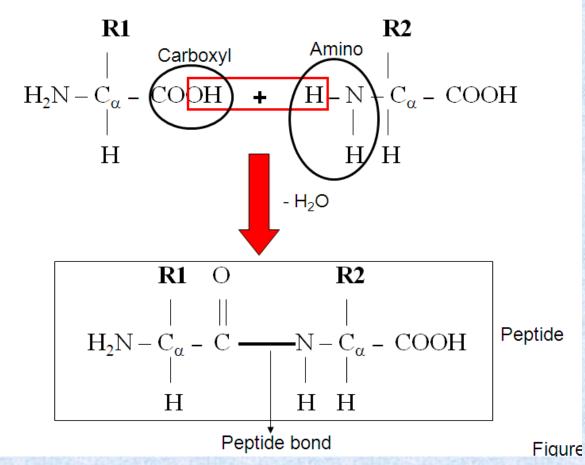
Amide > Acid > Alcohol > Amine > Ether > Alkane

Neutral (glutamine, asparagine) Acidic (glutamic acid, aspartic acid) Basic (lysine, arginine)

Polarity

- 1. Rank the following according to increasing non-polarity i.e. 1 = least non-polar, 4 = most non-polar; leu; phe; val; ala (A, V, L, F)
- 2. Rank the following amino acids by increasing polarity. i.e. 1 = more non-polar. ser; glu; asp; lys; ala; gln (A, K, S, E, D, Q)
- 3. Which amino acid is most insoluble in water: isoleucine or alanine? (I)
- 4. Which amino acid is most soluble in water: lys or ser? (S)

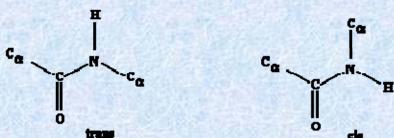
Peptide bond



Combination of amino acids

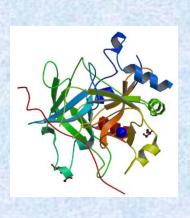
Elimination of water

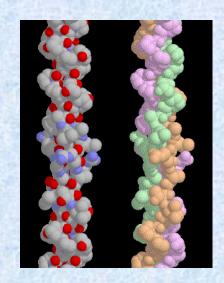
Peptide bond formation; strong

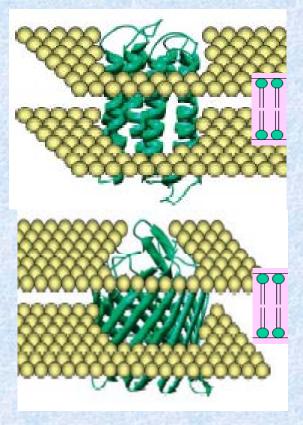


Proteins

Globular proteins
Fibrous proteins
Membrane proteins







Cytoplasm

Inner membrane α-helical

Periplasm

Outer membrane β-barrel (TMβ)

Outer space

Enzymes

Specialized proteins with catalytic activity (increase the rate of chemical reactions).

Only a small portion of the enzyme (3–4 amino acids) is directly involved in catalysis. The region that contains these catalytic residues, binds the substrate, and then carries out the reaction is known as the active site.

All the chemical reactions in cell are catalyzed by enzymes.

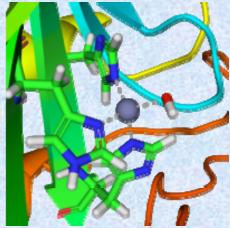
The **carbonic anhydrases** form a family of enzymes that catalyze the rapid interconversion of carbon dioxide and water to bicarbonate.

$$CO_2 + H_2O \xrightarrow{Carbonie\ anhydrase} H_2CO_3$$

(in tissues; high CO₂ concentration)

(in lungs; low CO₂ concentration)





three histidine residues and a hydroxide group coordinating (dashed lines) the zinc on at center.

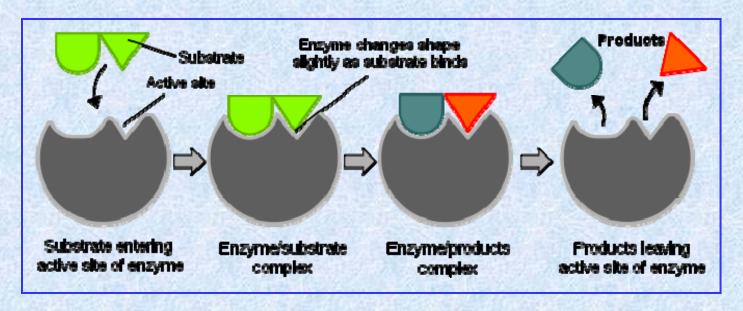
The active site of most carbonic anhydrases contains a zinc ion; they are therefore classified as metalloenzymes.

In enzymatic reactions, the **molecules** at the beginning of the process are called **substrates**, and they are converted into different molecules, called the **products**

(in tissues; high CO₂ concentration)

(in lungs; low CO₂ concentration)

Mechanism



Alcohol dehydrogenase: facilitate to break down alcohol (toxic) to others (aldehade/ketone) **Catalyzing oxidation**







The Comprehensive Enzyme Information System

	EC-Number	Enzyme Name Organism Pro Search Display 10 ▼ en	otein Full text Advanced Search ntries		
HIGH CHARLES OF BUILDING	Nevision & Preparation Nomenclature Enzyme Names EC Number Purification		Natural Substrate Products Natural Product Inhibitors Cofactors Metals/lons	Specific Activity pH Optimum pH Range Temperature Optimum Temperature Range Organism-related information	
SCHOOL IN THE PARTY.	Common/ Recommende Systematic Name Synonyms CAS Registry Number	Cloned Expression NEW Renatured Crystallization	Activating Compounds Ligands Biochemicals Reactions Aligned NEW	Organism Source Tissue Localization Protein-Specific Search	
		Stability	Enzyme Structure	Disease & References	
		pH Stability Temperature Stability	Sequence/ SwissProt link 3D-Structure/ PDB link	Disease/ Diagnostics References	Ε
		General Stability	Molecular Weight	Application & Engineering	
		Organic Solvent Stability Oxidation Stability	Subunits Posttranslational Modification	Engineering Application	

http://www.brenda-enzymes.org/

EC number: format

- Every enzyme code consists of the letters "EC" followed by four numbers separated by periods. Those numbers represent a progressively finer classification of the enzyme.
- For example, the tripeptide aminopeptidases have the code "EC 3.4.11.4", whose components indicate the following groups of enzymes:
- EC 3 enzymes are hydrolases (enzymes that use water to break up some other molecule)
- EC 3.4 are hydrolases that act on peptide bonds
- EC 3.4.11 are those hydrolases that cleave off the amino-terminal amino acid from a polypeptide
- EC 3.4.11.4 are those that cleave off the amino-terminal end from a tripeptide

Top level codes

			 [41]
Top-level	EC	num	bers

Group	Reaction catalyzed	Typical reaction	Enzyme example(s) with trivial name
EC 1 Oxidoreductases	To catalyze oxidation/reduction reactions; transfer of H and O atoms or electrons from one substance to another	$AH + B \rightarrow A + BH \text{ (reduced)}$ $A + O \rightarrow AO \text{ (oxidized)}$	Dehydrogenase, oxidase
EC 2 Transferases	Transfer of a functional group from one substance to another. The group may be methyl-, acyl-, amino- or phosphate group	$AB + C \rightarrow A + BC$	Transaminase, kinase
EC 3 Hydrolases	Formation of two products from a substrate by hydrolysis	AB + H ₂ O → AOH + BH	Lipase, amylase, peptidase
EC 4 Lyases	Non-hydrolytic addition or removal of groups from substrates. C- C, C-N, C-O or C-S bonds may be cleaved	RCOCOOH \rightarrow RCOH + CO ₂ or [X-A-B-Y] \rightarrow [A=B + X-Y]	Decarboxylase
EC 5 Isomerases	Intramolecule rearrangement, i.e. isomerization changes within a single molecule	AB → BA	Isomerase, mutase
EC 6 Ligases	Join together two molecules by synthesis of new C-O, C-S, C-N or C-C bonds with simultaneous breakdown of ATP	X + Y+ ATP → XY + ADP + Pi	Synthetase



BRENDA



The Comprehensive Enzyme Information System

EC 1.1.1.1 - alcohol dehydrogenase

Reaction catalyzed by alcohol dehydrogenase (1.1.1.1)

EC NUMBER COMMENTARY
1.1.1.1

RECOMMENDED NAME | GeneOntology No.

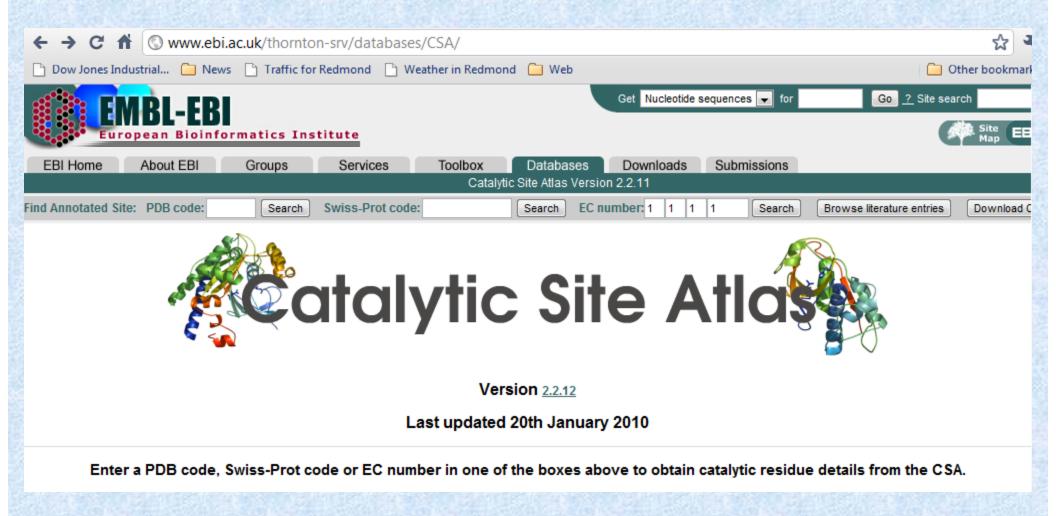
alcohol dehydrogenase GO:0004025

R
$$\rightarrow$$
 OH + NAD+ = \downarrow H + NADH + H+

A primary alcohol + NAD+ = an aldehyde + NADH + H+

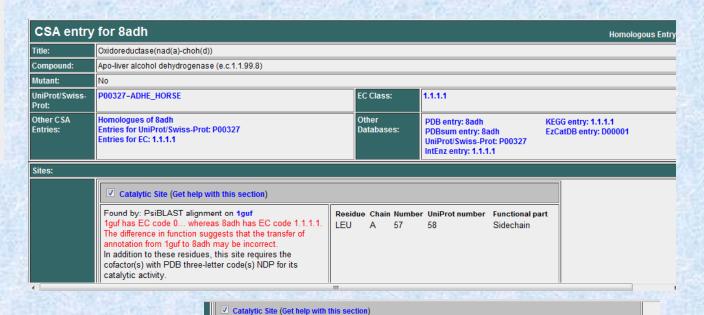
REACTION	REACTION DIAGRAM	COMMENTARY	ORGANISM	LITERATURE
A primary alcohol + NAD+ = an aldehyde + NADH + H+	<u>A</u>	ordered bi bi mechanism with cofactor adding first to form a binary enzyme complex	<u>Homo</u> <u>sapiens</u>	<u>285578</u>
A primary alcohol + NAD+ = an aldehyde + NADH + H+	<u>A</u>	Ser48 is involved in catalysis, isozyme gamma(2)gamma(2)	<u>Homo</u> <u>sapiens</u>	<u>654727</u>

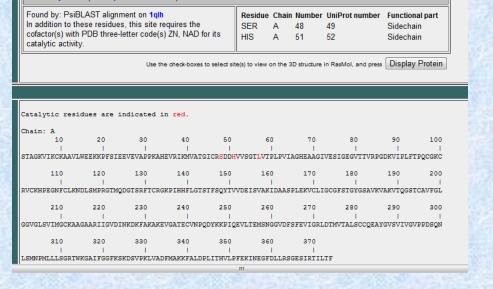
Catalytic site atlas



http://www.ebi.ac.uk/thornton-srv/databases/CSA/

EC: 1.1.1.1 **CSA Entries:** 1a4u, 1a71, 1a72, 1adb, 1adc, 1adh, ladf, 1adg, 1agn, laxe, laxg, 1b14, 1b15, 1b16, 1b21, 1bto, 1d1t, 1cdo, 1d1s, 1dda, 1deh, 1e3e, 1e3i, 1e31, 1ee2, lh2b, 1hdx, 1hdz, 1hdy, 1het, 1heu. 1hf3, 1hld, 1hso, 1hsz, 1ht0. 1htb. 1j5r, 1ju9, ljvb. 11de, 11dy, 111u, 1m6h, 1m6w, 1mc5, 1mg0, 1mgo, 1maO. 1mg5, 1n8k, 1n92, 1nto, 1nvg, 1qlj, 102d. lotq, 1p1r, 1qlh, 1qv6, 1qv7, 1r37, 1rjw, 1sby, 1teh. 1u3t, 1u3u, 1u3v, 1u3w, lvjo, 1ye3, 2adh, 2eer, 2fze, 2hcy, 2jhg, 2ohx, 2jhf, 2fzw, 3adh, 3bto, 3hud, 3i4c, 7adh, 5adh, 6adh, 8adh

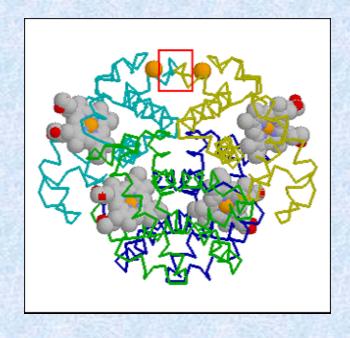


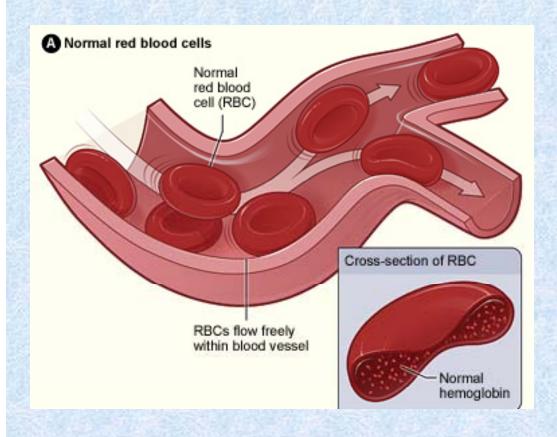


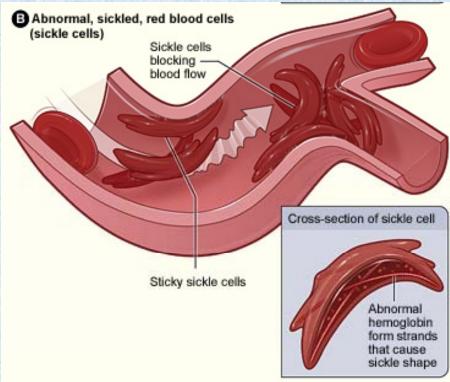
Transport proteins

Binds and carry specific molecules or ions from one organ to another

Hemoglobin is present in red blood cells, which efficiently carries oxygen from the lungs to the tissues of the body. Hemoglobin also helps in the transportation of carbon dioxide and hydrogen ions back to the lungs.







Defense proteins

Antibodies (immunoglobulins) can recognize and precipitate or neutralize invading bacteria or viruses from other species;

Thrombin is a **blood clotting protein** that prevent loss of blood when a vascular system is injured.

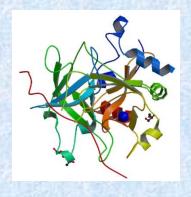
Regulatory proteins

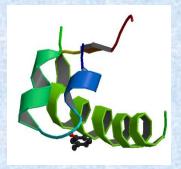
Regulate cellular or physiological activity.

Eg. insulin, which regulates sugar metabolism.

Others

Monellin has a intensely sweet taste, which is used to be a **sweetener**.







Nutrient and strorage proteins

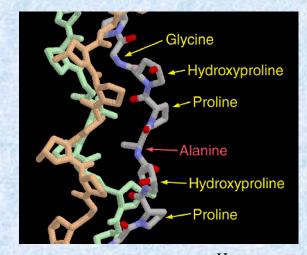
Ovalbumin, major protein of egg white, caesin in milk

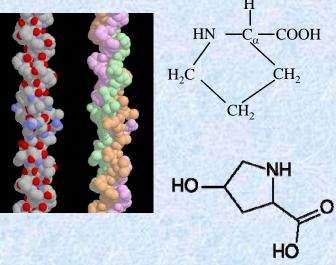
Contractile or Motile proteins

Actin and myosin function in the contractile system of skeletal muscle and many non-muscle cells.

Structural proteins

Give biological **strength/ protection**. Major component of tendons and cartilage is the fibrous protein, collagen. Hairs, finger nails and feathers consists of keratin.

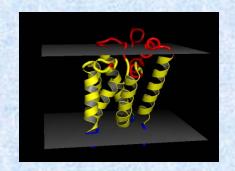


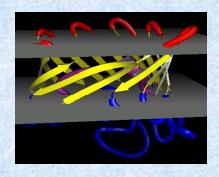


Hydroxyproline is produced by hydroxylation of the amino acid proline by the enzyme prolyl hydroxylase following protein synthesis

Membrane proteins

- Proteins that are embedded into membranes are called membrane proteins
- Structure: α-helical and β-barrel





• Function: Transporters, receptors, channels

Transporters

A membrane transport protein (transporter) is a membrane protein involved in the movement of ions, small molecules, or macromolecules such as another protein across a biological membrane.

The proteins may assist in the movement of substances by facilitated diffusion or <u>active transport</u>.

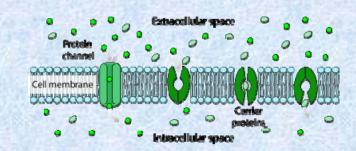
E.g. Multidrug efflux transporter AcrB

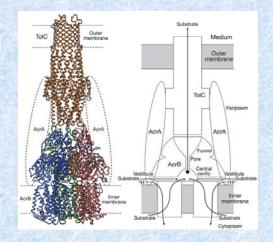
When substrate is transported AcR might recruit TolC to form direct transit pathway from the cytoplasm to extracellular.

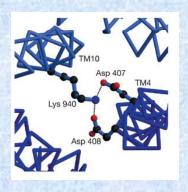
Two pathways (through cytoplasm; central cavity; central trimer hole plays a role in substrate translocation)

Ion pairs are possible candidates for transmembrane proton translocation site

Asp are protonated, ion pairs are disturbed; helices TM4 and TM10: conformation change -> transduced to the pore region by remote conformational coupling and open the pore.







Murakami et al. (2002) Nature 419, 587

Channels

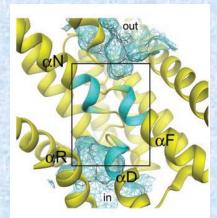
Ion channels are membrane proteins that selectively allow a given species of ion to pass through them.

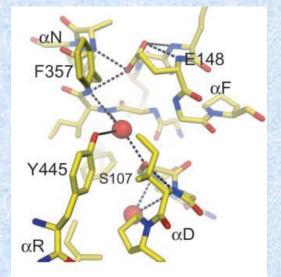
The two key properties are selective ion conduction and gating.

Selective conduction refers to a channel's ability to select one ion species among those present in the cellular environment.

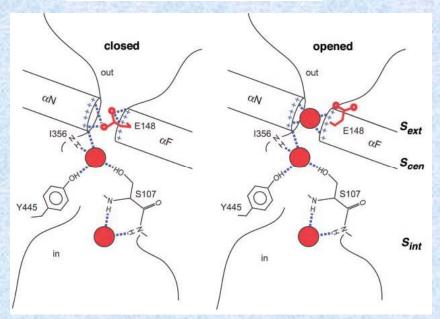
Gating refers to opening and closing the pore, the process by which ion conduction is turned on or off

Chloride channel
In: intracellular
Out: extracellular





Residues in the vicinity of chloride ions



Olfactory receptors

Olfaction is one of the senses involved in the perception of chemo-signals in the external environment.

The detection of odorant molecules involves specific binding to specialized receptor in olfactory system.

OR proteins belong to the G Protein-Coupled Receptor superfamily, which is characterized by the presence of hydrophobic transmembrane domains.

Each OR can recognize several chemically related molecules, and a specific odorant may bind to several ORs

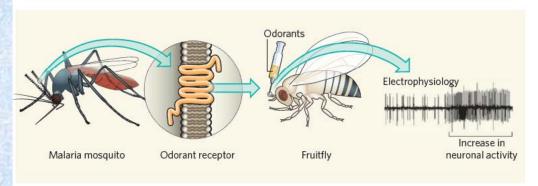


Figure 1| What's that smell? Odorant receptors (ORs) in the antennas of the malaria mosquito Anopheles gambiae detect odorants produced by humans. The genes that might encode ORs have been identified from the insect's genome⁴. Carlson and colleagues³ expressed 72 of these genes, one kind at a time, in the 'empty neurons' of the fruitfly Drosophila melanogaster. Of the putative ORs integrated into empty neurons, 50 turned out to be functional in the resulting mutant flies. The authors exposed these flies to a panel of 110 odorants, and measured the electrical activity of the OR-containing neurons. Odorants that bind to and activate ORs cause an increase in the number of spikes in the neuron's electrical activity. In this way, the authors identified which odorants activate (or, in some cases, inhibit) which ORs in A. gambiae.

Structure-Function Relationship

Protein structure

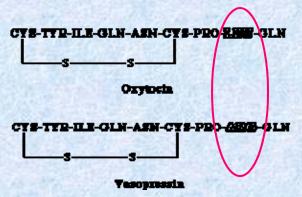


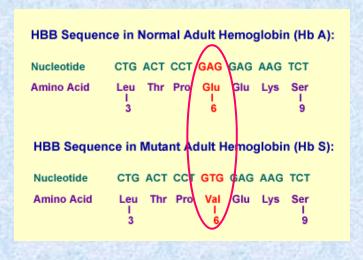
Protein function

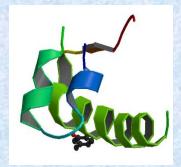
Function of a protein is dictated by its structure

Hormones

Simple change cause high blood pressure







MUTA	TION	PROTEIN	GENE_NAME	DISEASE
Phe 24	Ser INSULI	1	INS	DIABETES, MODY
His 10	Asp INSULI	1	INS	HYPERPROINSULINAEMIA
Phe 24	Leu INSULI	1	INS	HYPERPROINSULINAEMIA
Arg 39	His INSULI	1	INS	HYPERPROINSULINAEMIA
Arg 39	Leu INSULI	1	INS	HYPERPROINSULINAEMIA
Arg 39	Pro INSULI	1	INS	HYPERPROINSULINAEMIA
Val 42	Leu INSULI	1	INS	HYPERPROINSULINAEMIA

http://wiz2.pharm.wayne.edu/biochem/prot.html

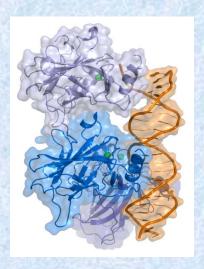
Structure-Function Relationship

Mutations in the p53 causes human cancer.

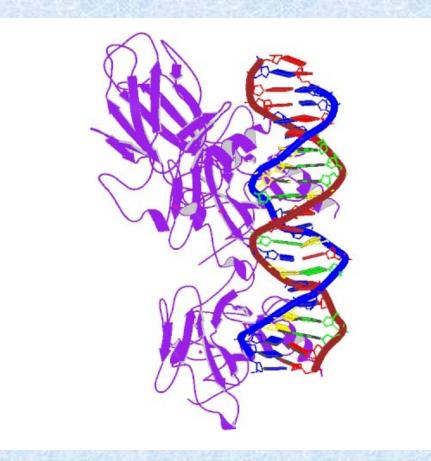
The majority of the mutations occur in the core domain (sequence-specific DNA binding activity of the p53 protein; residues 102-292)

Result in loss of DNA binding.

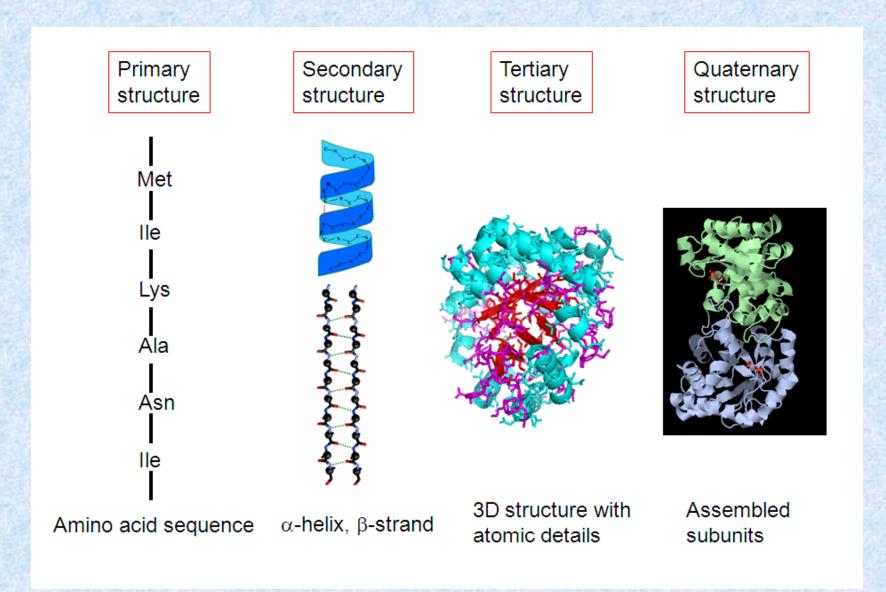
The structure supports the hypothesis that DNA binding is critical for the biological activity of p53



TUMOR SUPPRESSOR P53 COMPLEXED WITH DNA



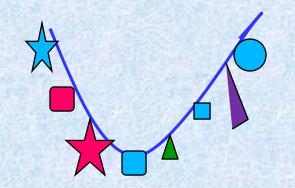
Protein Structure



Primary structure

It is a chain of amino acid residues in a specific order.

Chain with different types of beads





Nature selects a specific combination of amino acids to form a protein for its function.

26 alphabets:articles, books20 amino acidsproteins

Education: **OK**

University: **OK**

Edddddica: X

Uniiiiiivvvey: X

Primary structure: human hemoglobin

>sp|P68871|HBB_HUMAN Hemoglobin subunit beta OS=Homo sapiens

VHLTPEEKSAVTALWGKVNVDEVGGEALGRLLVVYPWTQRFFESFGDLSTPD AVMGNPKVKAHGKKVLGAFSDGLAHLDNLKGTFATLSELHCDKLHVDPENFR LLGNVLVCVLAHHFG KEFTPPVQAAYQKVVAGVANALAHKYH

Primary structure describes the linear sequence of amino acid residues in a protein.

It includes all covalent bonds between amino acids.

The relative arrangement of the linked amino acids is not specified.

Databases for protein sequences

EXProt

Munich Information Center for Protein Sequences (MIPS)

NCBI Protein database

PIR - Protein Information Resource (Georgetown University)

PIR-NREF

<u>PRF</u>

SWISS-PROT (Swiss Institute of Bioinformatics)

TrEMBL

UniProt - The Universal Protein Knowledgebase