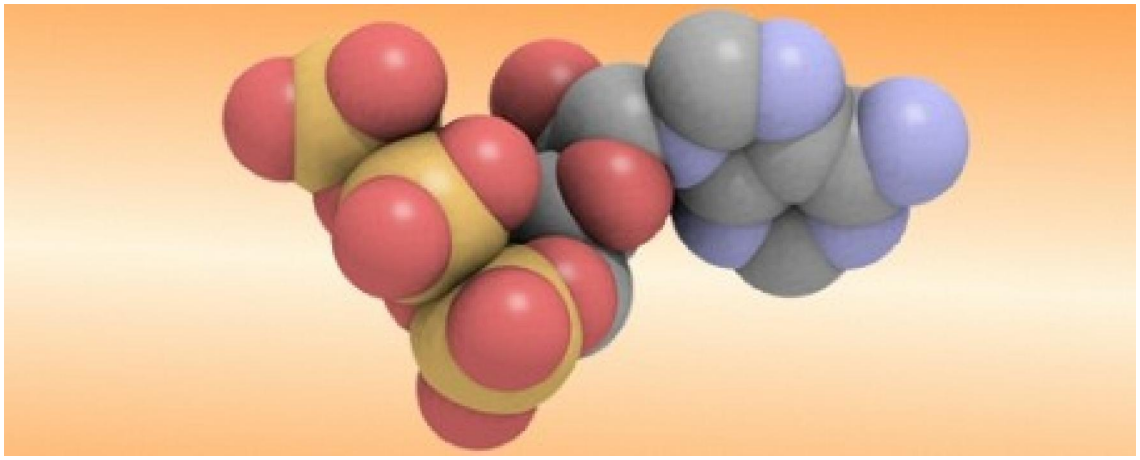
Biogeochemical cycles – Important Biomolecules

- The evolution of “synthetic” chemistry on the planet Earth is remarkable
- In particular, the ability to stabilise relatively weakly connected atoms with what we call chemical bonds has given a huge scope to the development of biology (life)
- Furthermore, relatively weak bonding forces can be used in concert to create remarkably robust structures, including many of the Earth’s creatures
- We are constantly learning from these examples and developing Chemistry correspondingly
- The complex nature of biological structures and their finely-tuned energetics can give amazing insights into the future possibilities for “Molecular Chemistry Evolution”

Some important biomolecules - ATP



Phosphorus is relatively rare on Earth, but is essential for life. The element P shows up in a surprisingly wide range of biological molecules. For instance, one of the best known molecules for carrying energy around our bodies is adenosine triphosphate (ATP).

Until recently, the leakage of phosphorus at all stages of the food production cycle was occurring with little fanfare, and phosphorus was more often than not labelled a pollutant for its effects on our waterways. Within the past five years, however, Australian-led research has sparked an international effort to raise awareness and foster sustainable management of this non-renewable resource which forms the basis of the global fertiliser industry.

Investigations by Dr Dana Cordell and Professor Stuart White from the Institute for Sustainable Futures at the University of Technology, Sydney predict that without action and at current rates the world will have consumed its best supplies of phosphorus within 20 years and may exhaust them by 2050.

http://eureka.australianmuseum.net.au/eureka-prize/environmental-research5

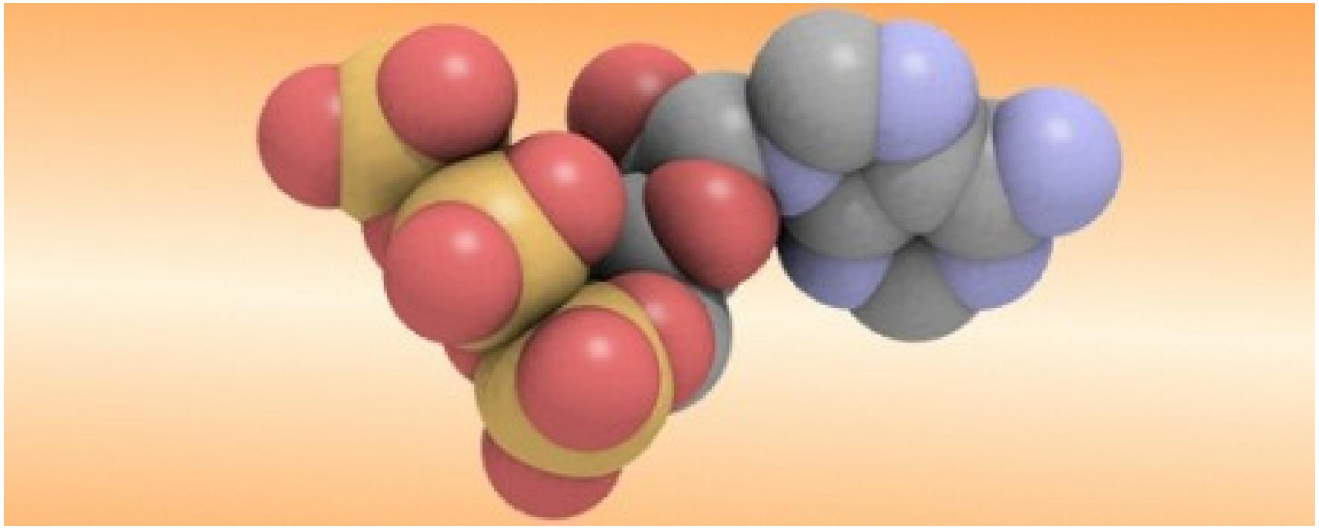
Environmental Research – Conserving life's building block

For their breakthrough work identifying phosphorus scarcity, tracking its life cycle and developing global and regional scenarios for its sustainable production and consumption, Dr Cordell and Professor White have been awarded the 2012 NSW Office of Environment and Heritage Eureka Prize for Environmental Research.

The Hon Robyn Parker MP, Professor Stuart White and Dr Dana Cordell
Photographer: Daniel O'Doherty

Some important biomolecules - ATP

- Metabolic processes that use ATP as an energy source convert it back into its precursors. ATP is therefore continuously recycled in organisms: the human body, which on average contains only 250 grams (8.8 oz) of ATP,[1] turns over its own body weight equivalent in ATP each day.

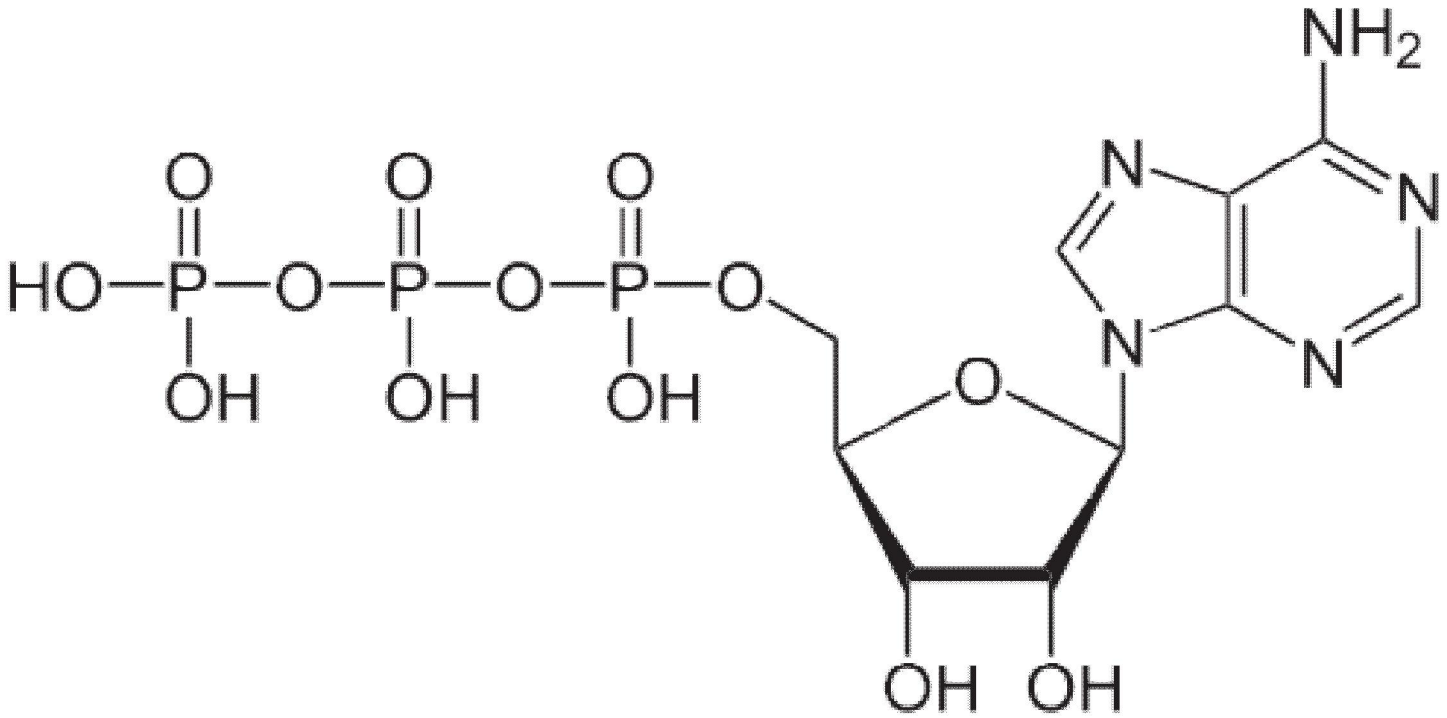


David E. Bryant, Katie E. R. Marriott, Stuart A. Macgregor, Colin Kilner, Matthew A. Pasek, Terence P. Kee. **On the prebiotic potential of reduced oxidation state phosphorus: the H-phosphinate-pyruvate system**. *Chemical Communications*, 2010; 46 (21): 3726 DOI: 10.1039/c002689a

Some important biomolecules - ATP

- All living things, plants and animals, require a continual supply of energy in order to function. The energy is used for all the processes which keep the organism alive.
- Some of these processes occur continually, such as the metabolism of foods, the synthesis of large, biologically important molecules, *e.g*. proteins and DNA, and the transport of molecules and ions throughout the organism.
- Other processes occur only at certain times, e.g. muscle contraction.
- Animals obtain their energy by oxidation of foods and plants by trapping sunlight using chlorophyll.
- Before the energy can be used it must be transformed into a form which the organism can handle easily. **This special carrier of energy is the molecule adenosine triphosphate, or ATP.**
- The ATP molecule is composed of three components.
- At the centre is a sugar molecule, ribose (same sugaras found in DNA).
- Attached to one side of this is the purine base adenine.
- The other side of the sugar is attached to a string of phosphate groups.
- The phosphates are the key to the activity of ATP.

Chemical structure of ATP



ATP consists of a base – far right - in this case adenine;
a ribose – middle
and a phosphate chain - left

How ATP works

ATP works by losing the endmost phosphate group when instructed to do so by an enzyme.

This reaction releases a lot of energy, which the organism can then use to build proteins, contact muscles, *etc*.

The reaction product is adenosine diphosphate (ADP), and the phosphate group either ends up as orthophosphate (HPO_4) or attached to another molecule (*e.g*. an alcohol).

$$ATP + H_2O \longrightarrow ADP + HPO_4$$

Even more energy can be extracted by removing a second phosphate group to produce adenosine monophosphate (AMP).

When the organism is resting and energy is not immediately needed, the reverse reaction takes place and the phosphate group is reattached to the molecule using energy obtained from food or sunlight.

Thus the ATP molecule acts as a chemical 'battery', storing energy when it is not needed, but able to release it instantly when the organism requires it.

Thermodynamic details of ATP reactions

- Any unstable system containing reactive molecules which are prevented from reacting is a means of storing free energy.
- In a cell this works by maintaining the concentration of reactive species far from the equilibrium point of the reaction.
- The standard amount of energy released from hydrolysis of ATP can be calculated from the changes in energy under standard conditions and then correcting to biological concentrations.
- The net change in heat energy (enthalpy) under standard conditions of the decomposition of ATP into hydrated ADP and hydrated inorganic phosphate is −20.5 kJ/mol.
- The free energy amounts released by cleaving either a phosphate (P_i) to give ADP or pyrophosphate (PP_i) to give AMP from ATP are
- $ATP + H_2O \rightarrow ADP + P_i \quad \Delta G^\circ = -30.5$ kJ/mol (−7.3 kcal/mol)
- $ATP + H_2O \rightarrow AMP + PP_i \quad \Delta G^\circ = -45.6$ kJ/mol (−10.9 kcal/mol)

Other important biomolecules

- Clearly there are many other biomolecules (and their individual building blocks) such as RNA, DNA, lipids, and so on with important functions within living systems
- In terms of biogeochemical cycles, proteins, and especially metalloproteins play vital roles in contributing to the chemical make-up of the environment
- Metalloproteins can fulfil a variety of functions
  - Catalysis – the field of metalloenzymes is huge
  - Small gas molecule transport and storage
  - Redox chemistry
  - Electron storage
  - Metal ion transport and storage
  - Small gas molecule fixation

Proteins – structural features

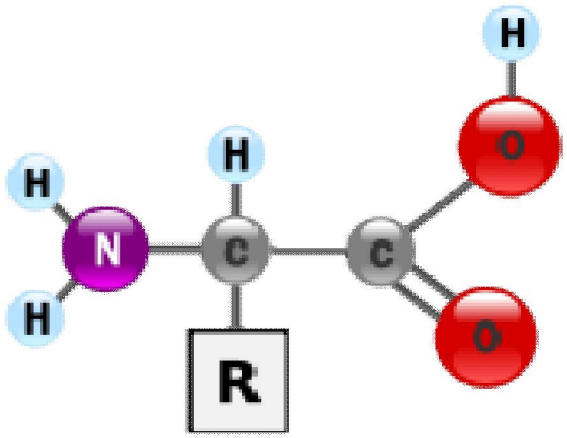
- There are 4 levels to the structure of proteins
- The PRIMARY level is the protein SEQUENCE – that is the order in which amino acids are strung together along the peptide backbone of the protein
- The SECONDARY level results from the unique nature of the peptide bond and hydrogen bonding interactions involving the backbone which can be intra- or interchain in nature and lead to the structural signatures of, amongst others, the α-helix and the β-sheet
- The TERTIARY level is provided by various “supramolecular” interactions involving the groups of the side-chains of the amino acids
- The QUATERNARY level involves interactions between protein subunits to give a complete protein system

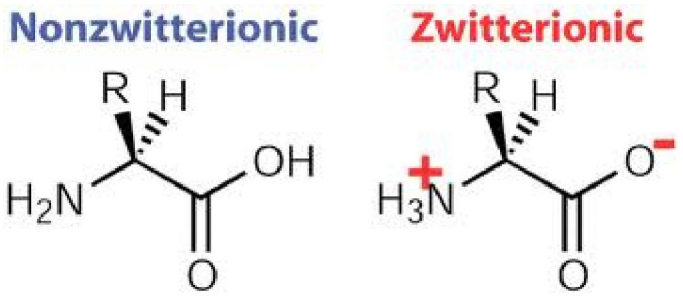
The PRIMARY level - the protein SEQUENCE

- Proteins contain specified arrangements of α-amino acids joined together via peptide bonds.
- There are 20 important naturally occurring amino acids and proteins contain 100s of amino acid units – so there are many possible combinations – but a given protein always shows the same sequence of amino acid side-chains.
- The macromolecule is formed via peptide bonds.
- Note that peptide bonds would not be expected by simply mixing amino acids together – rather salt/base chemistry should happen

The PRIMARY level - the amino acids

General form of an α-amino acid





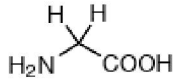
Amino acids tend to be in the zwitterionic form shown right

Peptides result from combining the carboxylate and amino groups of two amino acids with the formal elimination of a water molecule

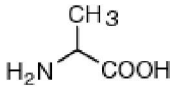
It is conventional to have the amino group left and the carboxylate group right – so-called N- and C- terminals

The PRIMARY level - the amino acids

Small

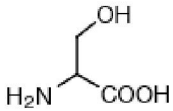


Glycine (Gly, G)
MW: 57.05

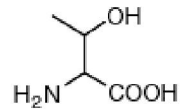


Alanine (Ala, A)
MW: 71.09

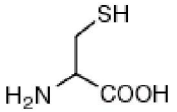
Nucleophilic



Serine (Ser, S)
MW: 87.08, pK_a ~ 16

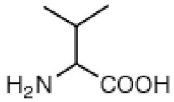


Threonine (Thr, T)
MW: 101.11, pK_a ~ 16

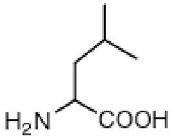


Cysteine (Cys, C)
MW: 103.15, pK_a = 8.35

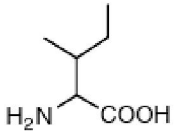
Hydrophobic



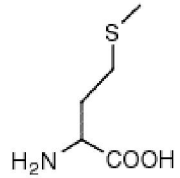
Valine (Val, V)
MW: 99.14



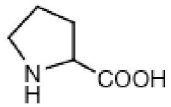
Leucine (Leu, L)
MW: 113.16



Isoleucine (Ile, I)
MW: 113.16

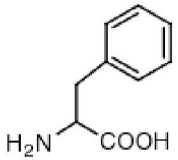


Methionine (Met, M)
MW: 131.19

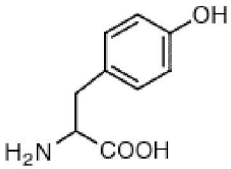


Proline (Pro, P)
MW: 97.12

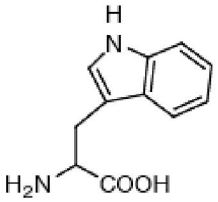
Aromatic



Phenylalanine (Phe, F)
MW: 147.18

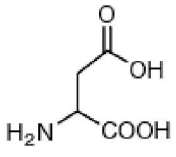


Tyrosine (Tyr, Y)
MW: 163.18

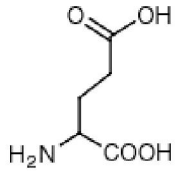


Tryptophan (Trp, W)
MW: 186.21

Acidic

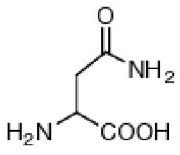


Aspartic Acid (Asp, D)
MW: 115.09, pK_a = 3.9

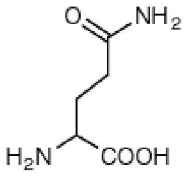


Glutamic Acid (Glu, E)
MW: 129.12, pK_a = 4.07

Amide

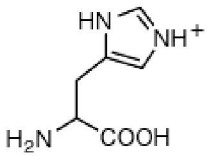


Asparagine (Asn, N)
MW: 114.11

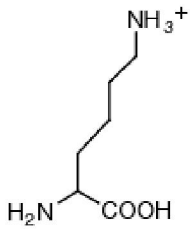


Glutamine (Gln, Q)
MW: 128.14

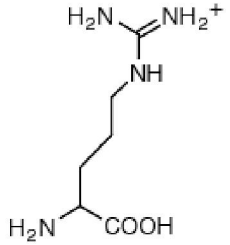
Basic



Histidine (His, H)
MW: 137.14, pK_a = 6.04

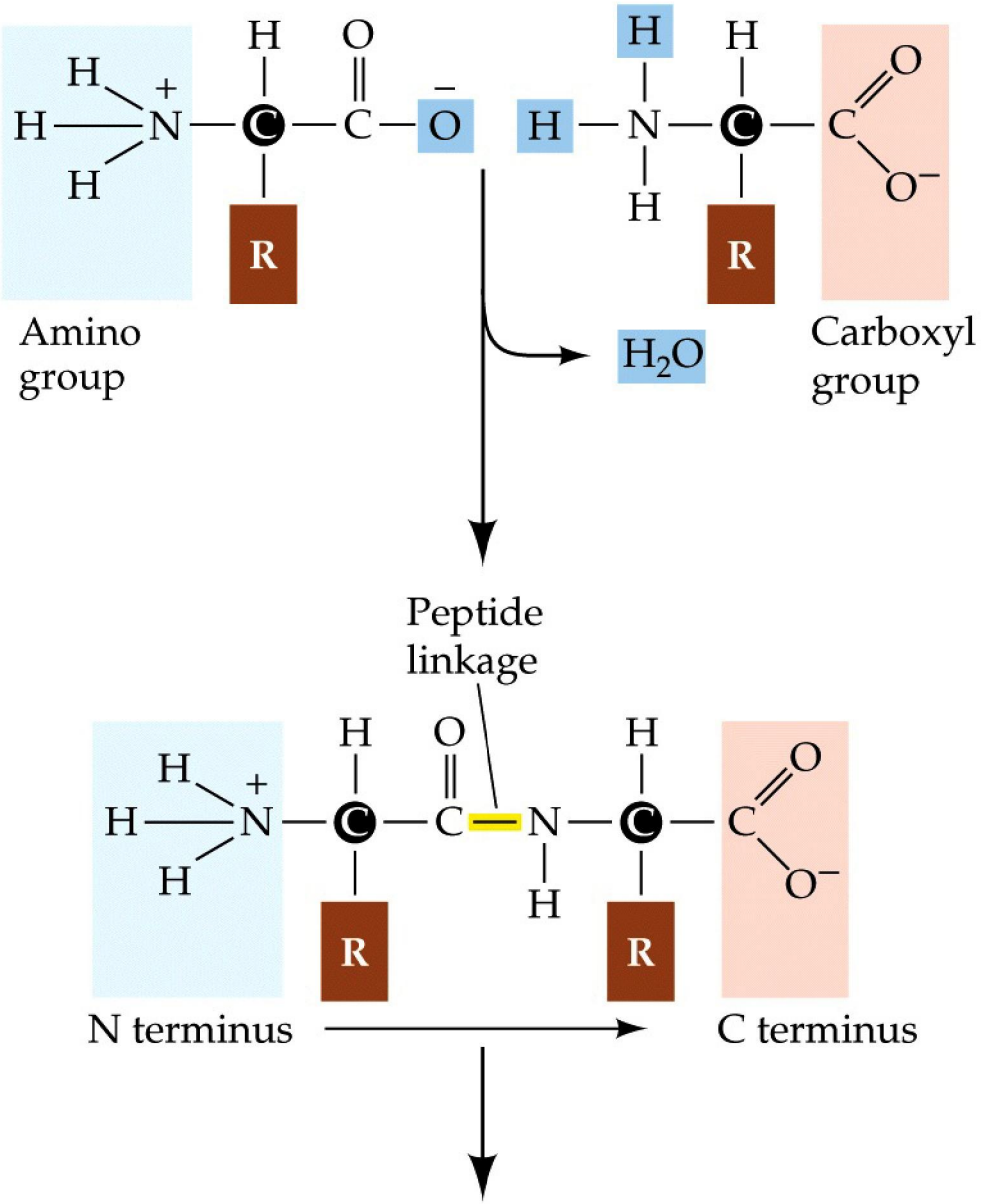


Lysine (Lys, K)
MW: 128.17, pK_a = 10.79



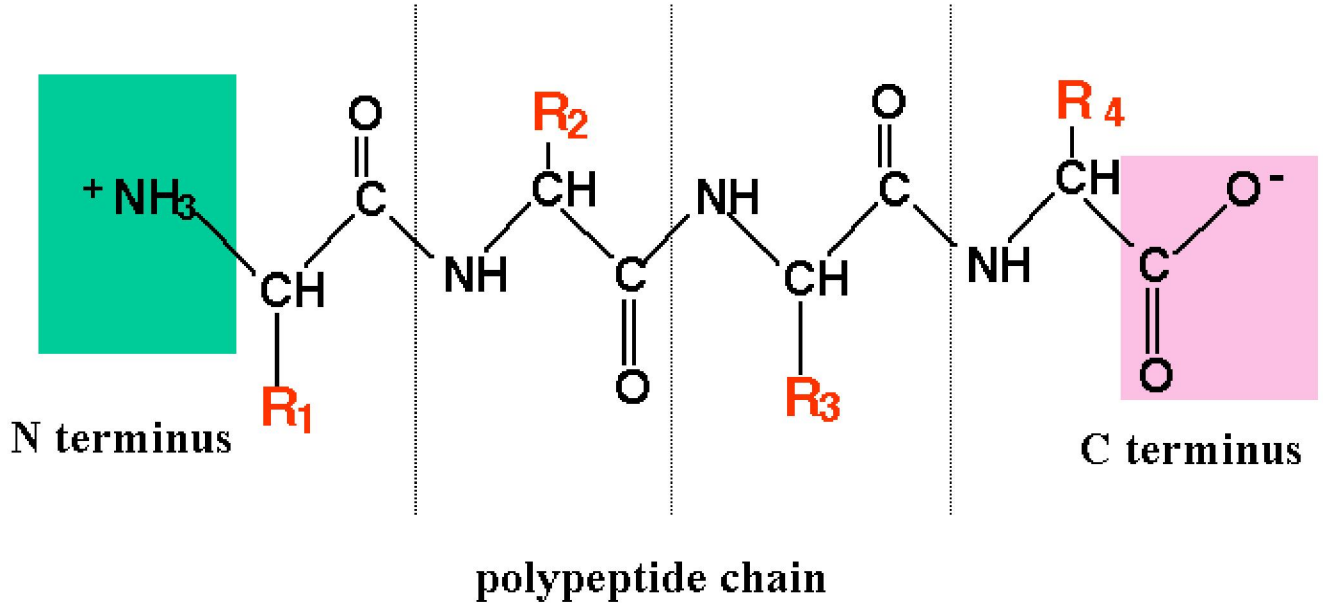
Arginine (Arg, R)
MW: 156.19, pK_a = 12.48

The PRIMARY level – the peptide bond



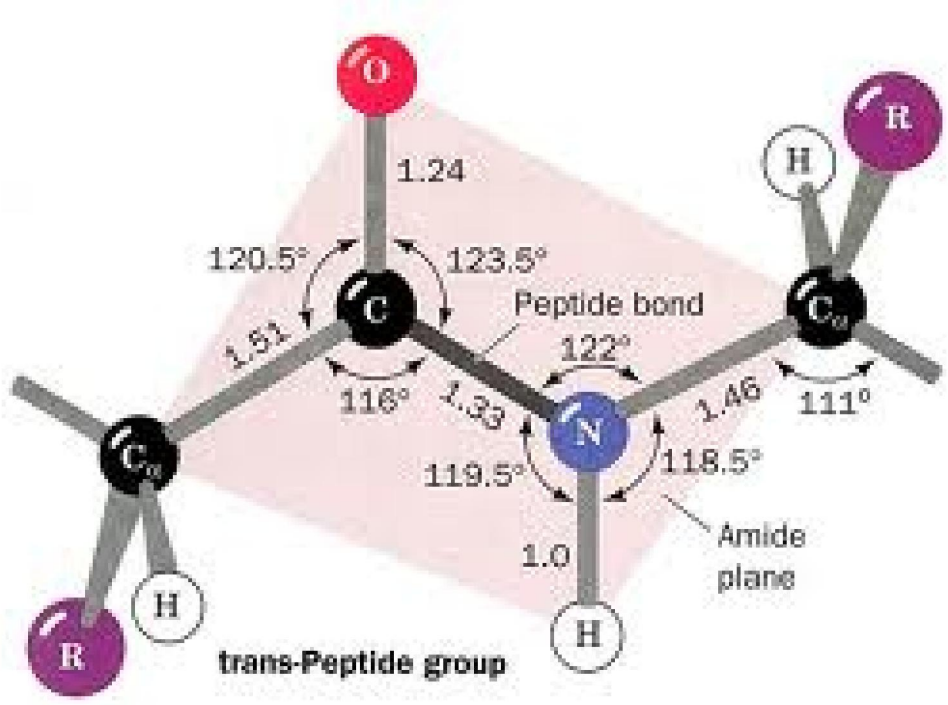
The PRIMARY level - the peptide bond

Peptide = chain of amino acids



The various R groups are important in terms of the final structure and function of the protein, but first the peptide bond itself influences the secondary structure.

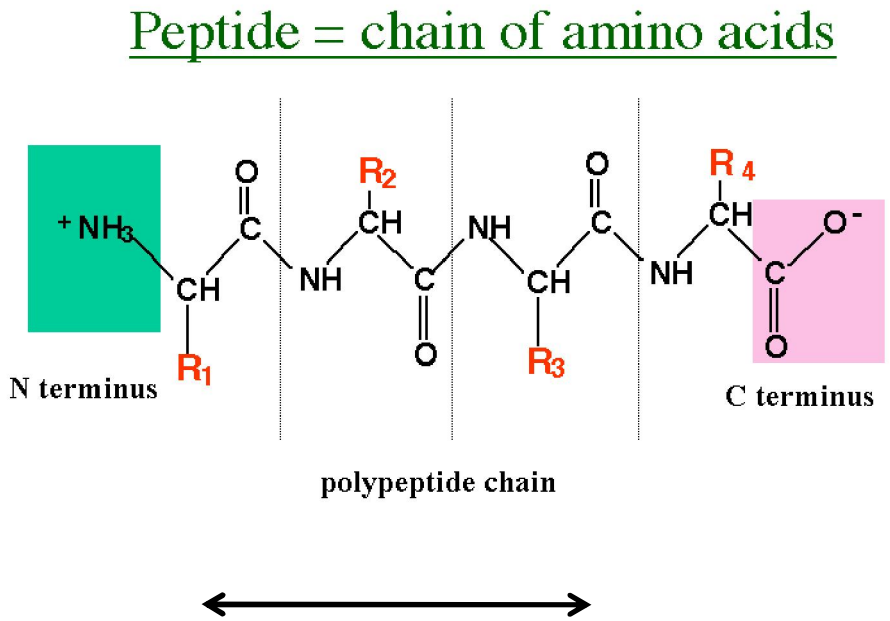
The SECONDARY level - the nature of the peptide bond



The C-N length in the peptide bond is shorter than expected whereas the C-O bond is longer, although the carbon is sp^2-hybridised. In fact, double-bond character between C and N leads to an essentially planar amide configuration.

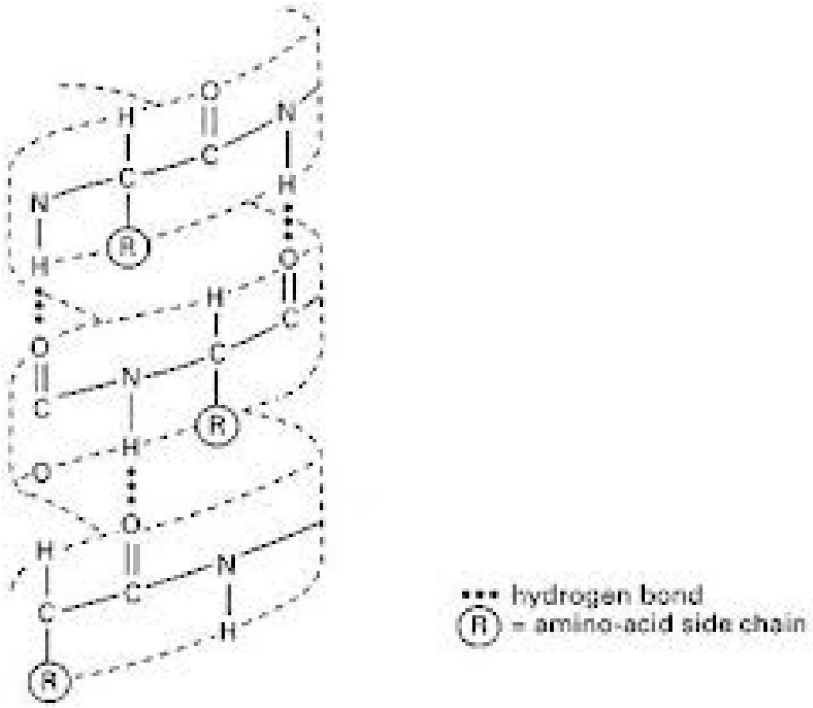
This can also be observed in more familiar amides such as DMA (dimethylacetamide) where the two N-bound amide methyl groups are found from NMR to be inequivalent as a result of restricted rotation about the C-N bond.

Secondary structure



Distance between residues (R1 and R3) on same side of backbone should be around 7.6 Å – actual distance found to be much smaller at ca. 5.4 Å

Secondary structure – the α-helix



The **alpha helix (α-helix)** is a right-handed coiled or spiral conformation, in which every backbone N-H group donates a hydrogen bond to the backbone carbonyl (C=O= group of the amino acid four residues earlier. This secondary structure is also sometimes called a classic **Pauling–Corey–Branson alpha helix**.

It is the most regular and easily identified structural motifs in proteins.

But how was this structure worked out?

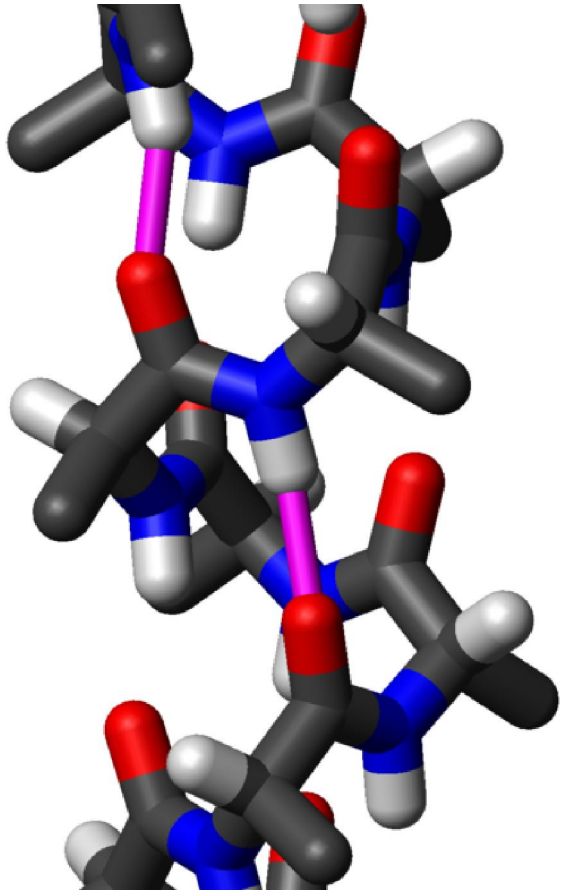
Secondary structure – the α-helix

- There were two key developments in modelling the α-helix structure:
- (1) the correct bond geometry, thanks to crystal structure determinations on amino acids and peptides leading to Pauling's prediction of *planar peptide bonds*
- (2) abandoning the assumption of an integral number of residues per turn of the helix.
- The turning point was in 1948, when Pauling, ill in bed with a cold, drew a polypeptide chain of roughly correct dimensions on a strip of paper and folded it into a helix, being careful to maintain the planar peptide bonds.
- After a few attempts, he produced a model with physically plausible hydrogen bonds.
- Pauling then worked with Corey and Branson to confirm his model.
- In 1954 Pauling was awarded his first Nobel Prize "for his research into the nature of the chemical bond and its application to the elucidation of the structure of complex substances" (such as proteins), prominently including the structure of the α-helix.

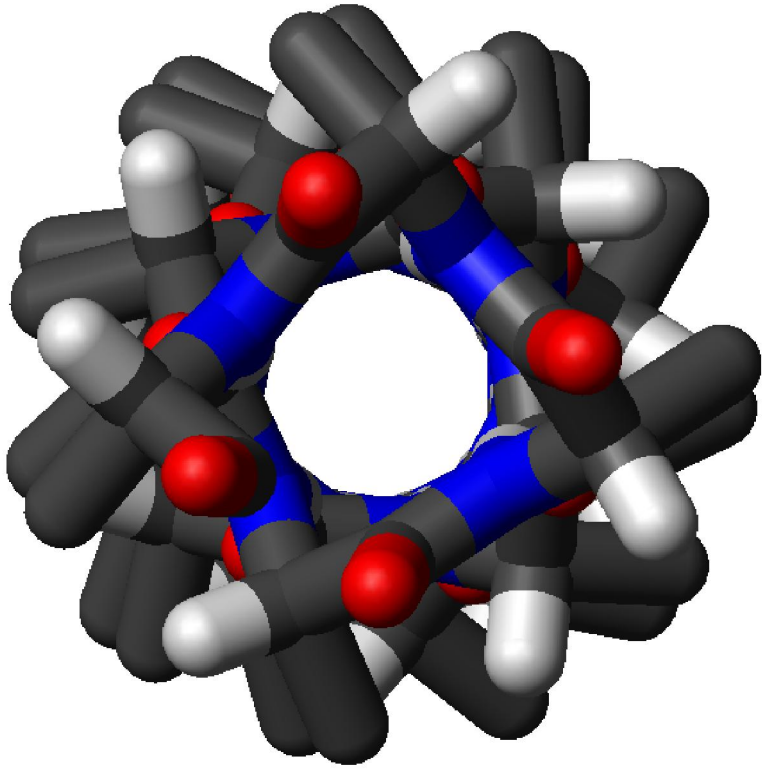
Secondary structure – Geometry and hydrogen bonding in the α-helix

- The following structural features are found in the α helix:
- Each amino acid residue corresponds to a 100° turn in the helix
- This implies that the helix has 3.6 residues per turn
- There is a translation of 1.5 Å along the helical axis
- The pitch of the alpha-helix (the vertical distance between one consecutive turn of the helix) is 5.4 Å (0.54 nm) which is the product of 1.5 and 3.6.
- This results from the hydrogen bonds between the N-H group from one amino acid with the carbonyl oxygen of the amino acid four residues ealrlier

Secondary structure – the α-helix

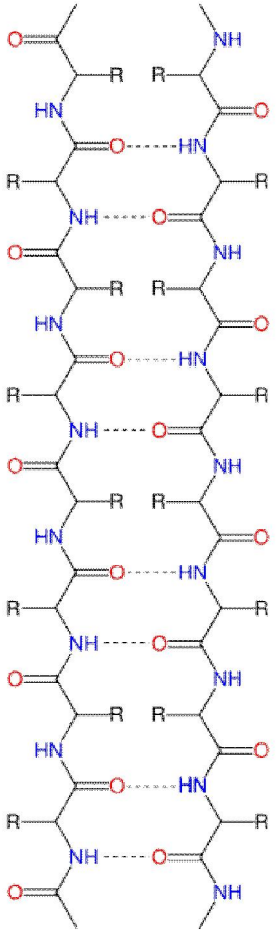


Side view of an α-helix of alanine. Two hydrogen bonds to the same peptide group are highlighted in magenta; the H to O distance is about 2 Å (0.20 nm)



Top view: four carbonyl groups point towards us spaced roughly 100° apart on the circle, corresponding to 3.6 amino acid residues per turn of the helix.

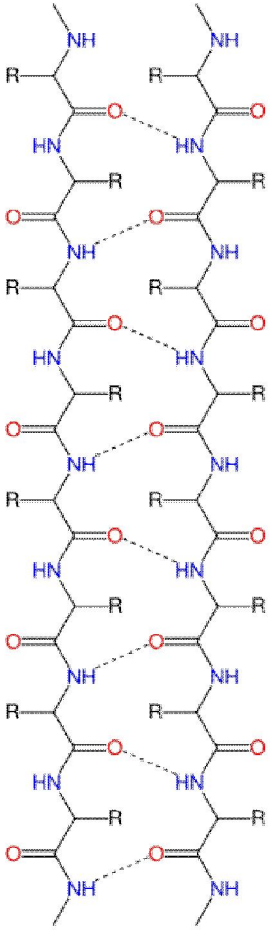
Secondary structure – the β-strand and sheet



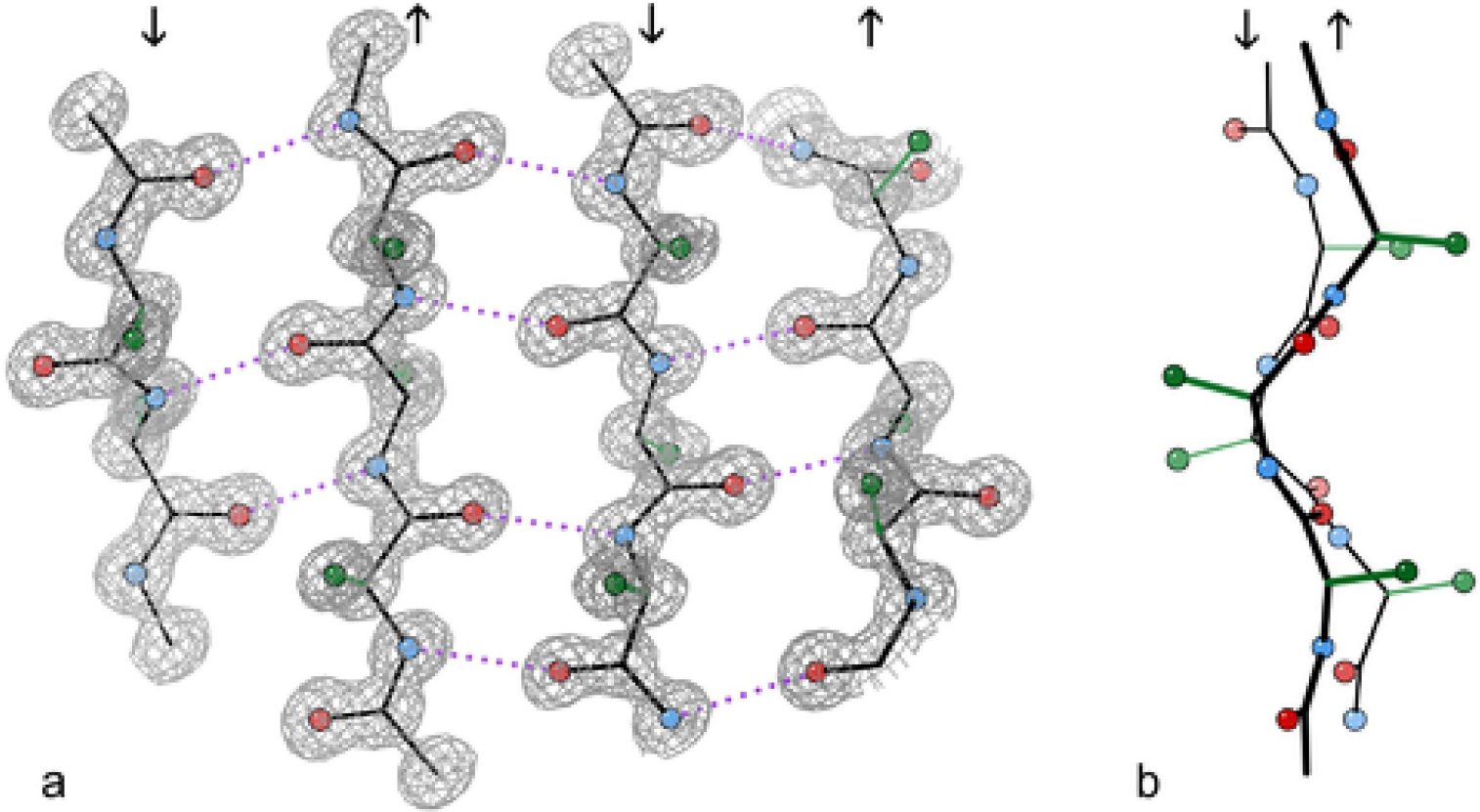
β strands are arranged adjacent to other strands and form extensive hydrogen bonds between chains (rather than along them). Collections of these give sheets.

The chains can be in antiparallel (left) or parallel (right) arrangements.

“Pleating” of the structure results from the local geometries at the carbon atoms.



Secondary structure – the β-strand and sheet



4-stranded antiparallel β sheet fragment from a crystal structure of the enzyme catalase.

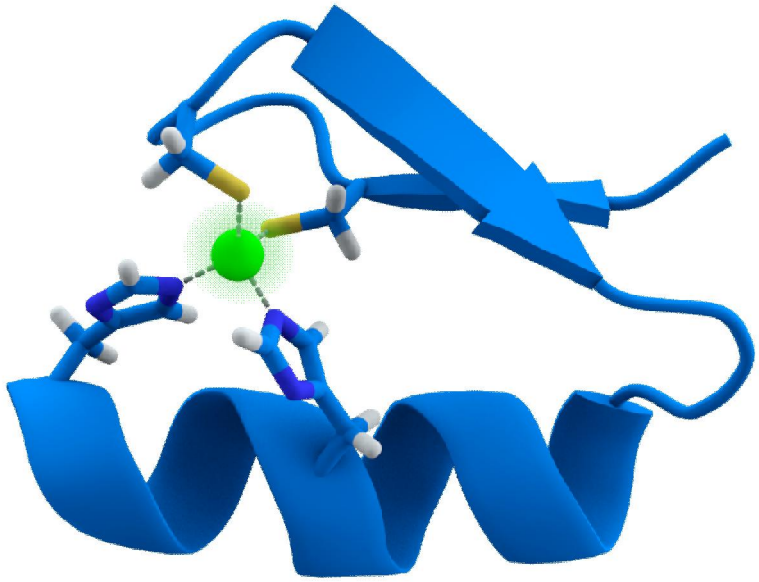
a) Front view, showing the antiparallel hydrogen bonds (dotted) between peptide NH and CO groups on adjacent strands. Arrows indicate chain direction, and electron density contours outline the non-H atoms.

b) Edge-on view of the central two β strands in a, showing the right-handed twist and the pleat motifs.

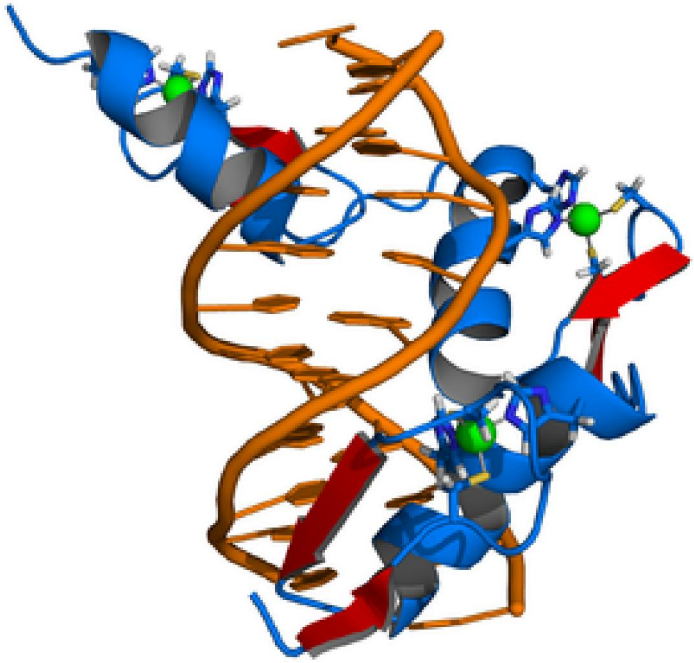
The TERTIARY level

- The TERTIARY level is provided by various “supramolecular” interactions involving the groups of the side-chains of the amino acids, for example:
- Hydrogen bonding interactions between carbonyl groups and other protic groups such as alcohol groups on serine carboxylate groups on aspartate
- Hydrophobic interactions (dipolar and quadrupolar) between bulky organic groups such a phenyl in phenylalanine and tryptophan bezene groups
- Electrostatic (i.e. to some extent ionic) interactions though formation of ion pairs (carboxylate with ammonium groups)
- Cross-coupling through formation of disulfide bonds from thiolate residues on e.g. cysteine
- Structural metal ion features such as the coordination bonds formed by Zn(II) to stabilise superoxidedismutase, alcohol dehydrogenase and zinc finger protein motifs

The TERTIARY level – Zn-Fingers



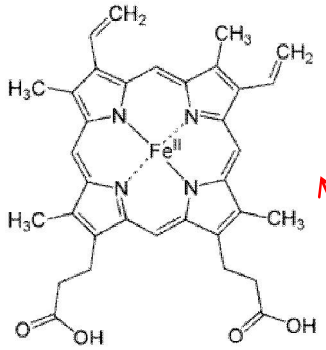
Cartoon of the Cys2His2 zinc finger $\{Zn(N(His)_2S(Cys)_2)\}$ motif combining α- helix and antiparallel β-sheet secondary structural elements



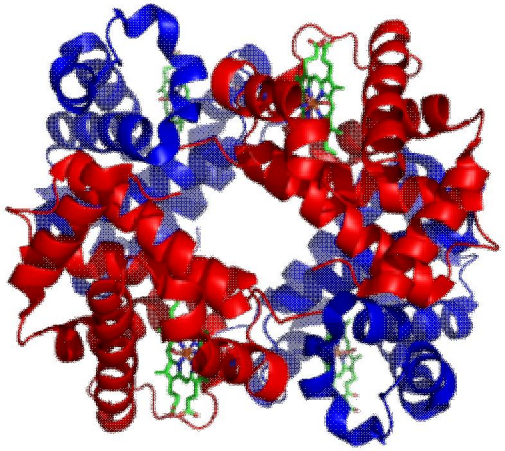
Cartoon representation of the Zn-finger protein zif268 (blue) interacting with DNA (orange). The three Zn(II) centres are in green.

The QUATERNARY level

This involves interactions between protein subunits to give a complete protein system



+ protein then x4



Haemoglobin transports oxygen in blood using a porphyrin bound Fe(II) ion to capture the molecule. The high efficiency of the system is the result of cooperative binding effects arising from the concert of 4 protein subunits (one Fe(II) per subunit) in which subtle structural changes on oxygen binding to one Fe(II) tip the system towards further oxygen binding.