Inorganic Chemistry of Life
Chemistry 2211a

Instructor: Martin Stillman ChB064
Martin.stillman@uwo.ca

3) Biology and Biochemistry important for bioinorganic chemistry

Amino acids
pK_a, pH, pI – amino acids
Protein structure
Nucleic acids – DNA and special structures
Making proteins
Cells
Membranes

(Check with "Late Breaking News" on URL instruct.uwo.ca/chemistry/2211a for changes.)

Recommended text Books

Principles of Bioinorganic chemistry by Lippard & Berg. TAYSTK QU 130.L765 1994 (On heavy demand (2-hour loan) at the Taylor Library and in the book store.)

Bioinorganic chemistry: a short course by Roat-Malone. QU130.R628b (On heavy demand (2-hour loan) at the Taylor Library and in the book store.)

Bioinorganic chemistry: inorganic elements in the chemistry of life: an introduction and guide by Kaim and Schwederski. (On heavy demand (2-hour loan) at the Taylor Library.)

The biological chemistry of the elements: the inorganic chemistry of life by da Silva and Williams. QU4.S586b 2001 (On heavy demand (1-day loan) at the Taylor Library.)

There are many biochemistry books in the library that you could consult for background.

2211a-3-2016-BIO-L15-r16-f10.doc r16-f10

To start then

Summary: This section provides the background necessary to understand how biological chemistry allows us to interpret the interaction of metals with the different molecules found in biological systems.

1) The polymerisation of amino acids into proteins – really at first just peptide chains – the formation of the 3D structures, in some cases requiring a metal (- see Zn in zinc finger proteins) or needing metal cofactors to function (Zn, Mg, heme, FeS clusters), activates the protein. Structure controls function. (See L&B p 51 for an example)

2) The formation of ribonucleic acids (RNA) and deoxyribonucleic acids (DNA) from nucleotides - polymerization is the basis of life.

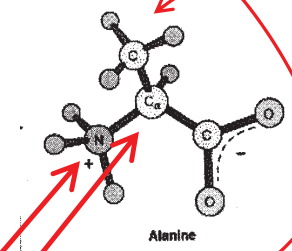
1. While for many decades the helical DNA was thought to be essentially inert with just the genetic code as its role,
2. it is now known that DNA takes part in many reactions
3. and that many different enzymes interact specifically with the coiled DNA
4. making use of the minor and major groove differences (see the Intro unit and the Zn unit and L&B p 183 for an example of a Zn-finger protein binding to DNA, the barrels of α helix lie in the grooves of the double helix)
5. electrical conductivity along the length of the double helix has been suggested as a means of biological messaging
6. indeed, the double helix does conduct electricity under certain circumstances!

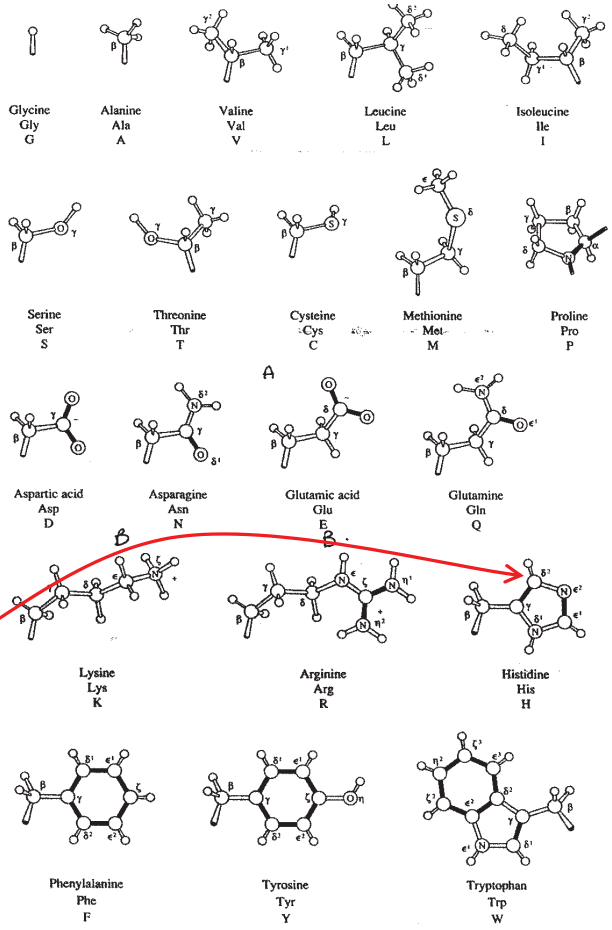
L-B	R-M	K-S	Problems to do
Ch 3 – p 43-74	Ch 2 p 24-48; 57-65		If blank – see later

3083

1. Amino acids

1. 20 natural amino acids (predominate)
2. 10 must be part of the diet – Arg, His, Ile, Leu, Lys, Met, Phe, Thr, Trp, Val
3. in proteins all are L-stereoisomers
4. but other amino acids (eg L-homoserine; L-selenocysteine) and D-isomers are found in other compounds, eg cell walls, even as neurotransmitters in the brain – sugars are all D in DNA and RNA – so why L-isomers only? Not known.
5. at pH 7 exist as the zwitter ion, ie -COO^- and -NH_3^+
6. the terminology for naming involves the carbon on main chain – it is alpha – connecting to the first atom attached to this carbon being beta atom. Then gamma, delta, ε,
7. with rings, the naming takes place on both sides – so we find δ1 and δ2 in His, etc. – we see a bit more detail on the next slide

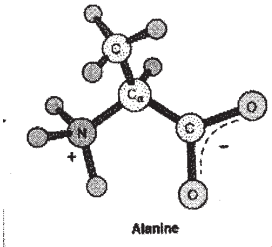


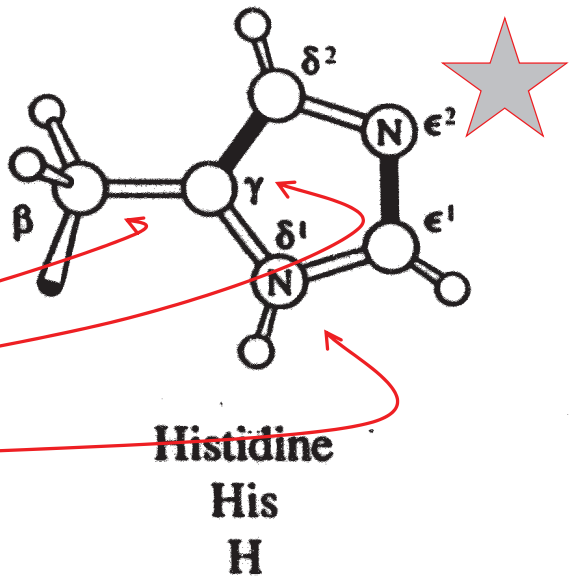


3083

1. Amino acids

1. 20 natural amino acids
2. 10 must be part of the diet – Arg, His,Ile, Leu, Lys, Met, Phe, Thr, Trp, Val
3. in proteins all are L-stereoisomers
4. but other amino acids (eg L-homoserine; L-selenocysteine) and D-isomers are found in other compounds, eg cell walls, even as neurotransmitters in the brain – sugars are all D in DNA and RNA – so why L-isomers only? Not known.
5. at pH 7 exist as the zwitter ion, ie -COO^- and -NH_3^+
6. the terminology for naming involves the carbon on main chain – it is alpha – connecting to the first atom attached to this carbon being the beta atom. Then gamma, delta, ε, etc.
7. with rings, the naming takes place on both sides – so we find δ1 and δ2 in His, etc.





Key metal binding amino acids are:

a. His – delta N or epsilon N. (δ or ε)
b. Asp – with the possibilities we saw earlier for different bonding modes through the $-COO^-$ group.
c. Cys
d. Met
e. Glu (note 1 CH_2 extra in Glu c.f. Asp)
f. Tyr

You will also find the amino acids referred to by their single letter in sequence tables for proteins.

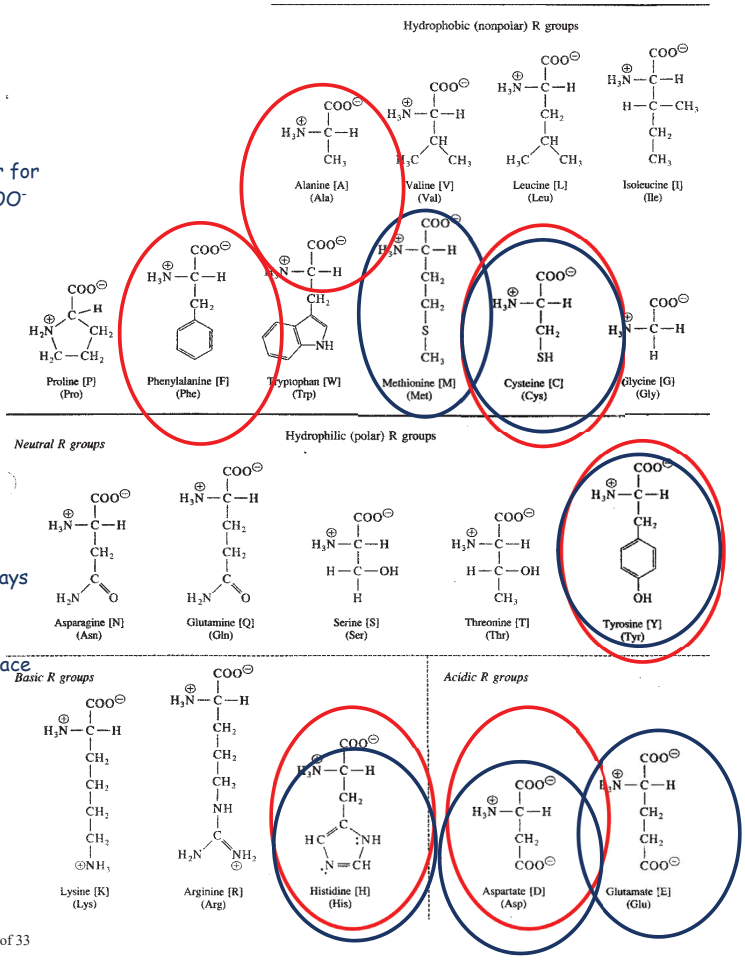
(Must be able to draw and recognise (a) – (f))

Classes of amino acid shown here:

g. Hydrophobic – inc aromatic
h. Hydrophilic – Tyr due to the –OH
i. Basic – strongly polar, side chains are always positively charged at neutral pH
j. Acidic – side chains are negative at pH 7; hydrophilic because so polar; on the surface of proteins to solubilize the protein

Good examples to memorize because …

often bind to metals



Essential	Nonessential
Isoleucine	Alanine
Leucine	Arginine*
Lysine	Aspartate
Methionine	Cysteine*
Phenylalanine	Glutamate
Threonine	Glutamine*
Tryptophan	Glycine*
Valine	Proline*
Histidine	Serine
Asparagine* - in disease	Tyrosine*
Selenocysteine** May be required in the diet	

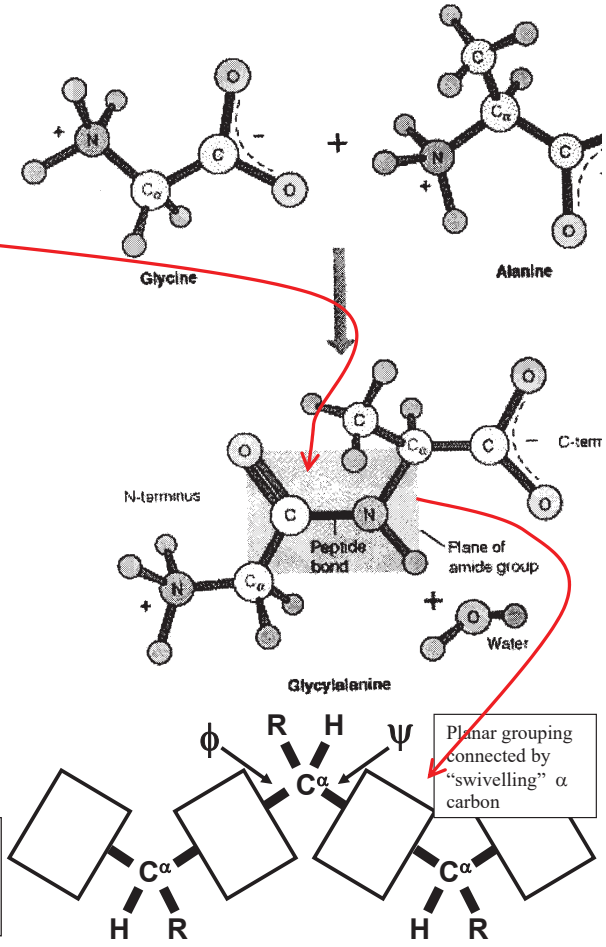
What to know from this section..

1. 20 amino acids – many combinations possible in a peptide chain → a protein if >20 'residues', polypeptide if smaller – that's how function is built up - don't need to memorize all 20 – need to know "20"
2. 10 in diet – don't need to memorize all 10 – but need to know "10"
3. Structure of the basic amino acid – see glycine – draw it
4. Structure of amino acids with side-chains and rings – label the carbons
5. Be able to draw and recognise Gly – Ala - Cys – Met- -Asp – Glu – His – Tyr
6. What makes an amino acid basic?
7. What makes an amino acid acidic?
8. Recognise hydrophobic amino acids – aromatic
9. Recognise hydrophilic amino acids – all have –OH group
10. Most important – be able to describe the key metal-binding side chains – that is the key metal-binding AAs
11. Be able to draw, list – recognise - the key metal-binding amino acids
12. Associate the class of amino acid with the metal-binding set in (6)

3124a

Formation of proteins

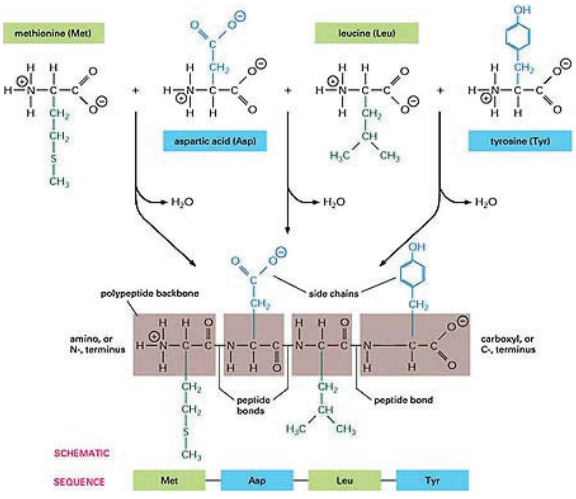
1. First, the peptide chain, forms through elimination of water in a condensation reaction between the COO^- and $-NH_3^+$ groups.
2. The peptide bond has the special property of being planar due to overlap of the CO and NH bonds through space, this reduces the reactivity of the CO and –NH- groups and maintains a rigid localized structure.
3. The primary structure refers to the assembly of amino acids into the sequence that defines the protein.
4. The N-terminal is the beginning of the chain; the amino acid is #1. It is the first synthesized (see next slide).
5. The secondary structure refers to the local conformation extending over a few amino acids often determined by hydrogen bonding between electropositive hydrogens and electronegative O, N, F, Cl.
   Special structures form:
   helix coiling
   β pleated sheets both parallel and antiparallel
6. The tertiary structure defines how these locally organized regions fold together
7. The quaternary structure defines how different polypeptide chains assemble into a single more massive unit.



L-B	R-M	K-S	Problems to do
44-47 48 for 2^0	28-29 30 for 2^0		

Partial double bond - -C=O-NH- to $-CO^-=NH-$ this leads to rigid groupings between amino acids Phi and Theta can change to allow 3D structures to form – as long as the side chains don't collide!

Example of peptide chain – study point – write this sequence out in the correct order using single letter codes.



Which of the amino acids here could bind to a metal?

Give an example of a protein with each of these metal-binding amino acids used to bind a metal:

Write out a modified sequence as 3-letter codes and by drawing the molecules so that Zn^{2+} would be the likely metal in a solely structural role.

Protein Sequences – (7) primary structure – bioinformatics –

http://en.wikipedia.org/wiki/Sequence_alignment_software

-examining the primary stucture for clues to

(i) metal binding (how?)

(ii) structure – regions with known structural proerties - we can recognises AAs for turns, sheets, helix

(iii) evolutionary changes – point mutations – single amino acid switches

(iv) key programmes – BLAST "http://blast.ncbi.nlm.nih.gov/Blast.cgi"

"The Basic Local Alignment Search Tool (BLAST) finds regions of local similarity between sequences. The program compares nucleotide or protein sequences to sequence databases and calculates the statistical significance of matches. BLAST can be used to infer functional and evolutionary relationships between sequences as well as help identify members of gene families."

(v) conserved regions – that is regions of the sequence in proteins from different organs or diferent species, can provide informaiton on specific metal binding, structural or functional signfiicance. Example from above url – 2 Zn-finger proteins –

```
AAB24882        TYHMCQFHCRYVNNHSGEKLYECNERSKAFSCPSHLQCHKRRQIGEKTHEHNQCGKAFPT 60
AAB24881        --------------------YECNQCGKAFAQHSSLKCHYRTHIGEKPYECNQCGKAFSK 40
                                    ****: .***:  * *:** * :****.:* *******..

AAB24882        PSHLQYHERTHTGEKPYECHQCGQAFKKCSLLQRHKRTHTGEKPYE-CNQCGKAFAQ- 116
AAB24881        HSHLQCHKRTHTGEKPYECNQCGKAFSQHGLLQRHKRTHTGEKPYMNVINMVKPLHNS 98
                 **** *:***********:***:**.: .***************     :  *.: :
```

A sequence alignment, produced by ClustalW, of two human zinc finger proteins, identified on the left by GenBank accession number. Key: Single letters: amino acids (see 3 slides up for decoding – D, E, M, Y, and most important **H & C**).. Red: small, hydrophobic, aromatic, not Y. Blue: acidic. Magenta: basic. Green: hydroxyl, amine, amide, basic. Gray: others. "*": identical. ":": conserved substitutions (same colour group). ".": semi-conserved substitution (similar shapes).

An excellent page is here: http://en.wikipedia.org/wiki/Amino_acid– **STUDY: list all metal-binding AAS**

Protein Sequences – (8) secondary structure

Stabilization energies in biological systems – this is what holds everything together in solutions…(myoglobin – 8 helical 2^0 segments and 3^0 structure)

Weak -purely ionic interactions +..- between oppositely charged atoms
Weak - dipole interactions here the dipole moments from the C=0 and N-H
Hydrogen bonds - much stronger because so many
Van der Waals or dispersion forces – atoms close to each other cause polarization of the electrons – but only over short distances – too close and repulsion
Hydrophobic interactions – bundling the peptide chain to keep water away but maximize the H-bonding on the outside in the hydrophilic region

(i)	Chemical bonds	C-H	105	kcal.mol^{-1}
(ii)		C=C	172	
(iii)	Ionic hydration	Na^+ -	93	Ca^{2+} - 373
(iv)	Hydrogen bonds	O…H -	5	(in vacuum)
(v)	Protein folding	~ 2-10	(in solution)	

(vi) The secondary structure refers to the local organization of conformations extending over a few amino acids often determined by hydrogen bonding between electropositive hydrogens and electronegative O, N, F, Cl. (Always H in one bond and e-neg atom in another.)

(vii) Special structures form: helix coiling pleated sheets both parallel and antiparallel

Because the amino acid sequence (1^0 structure) is determined by the 'random' selection from 20 amino acids to form chains of any length – there are many permutations – so no duplication of sequences … except…

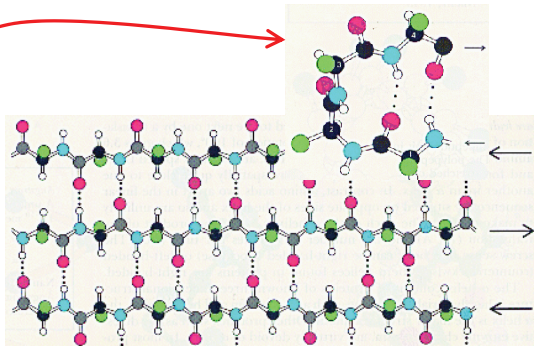
We want structures (2^0 structure):

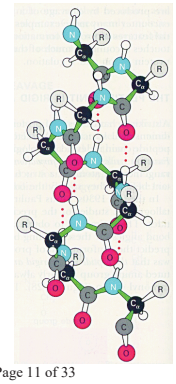
– so helical structures (right-handed α helix is most common – left-hand has been reported) are set up by certain amino acid groupings

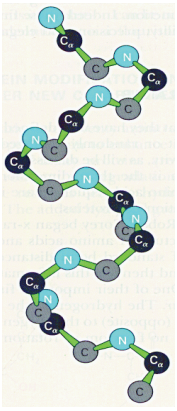
- β pleated sheets are triggered by a β turn

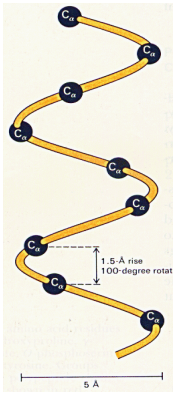
- arrows N → C ie from N terminus to the C terminus

All held together by hydrogen bonds between the peptide bond – so the side chains are not essential – need to be small so as not to disrupt the strands too much.









1058

Examples of β pleated sheets – parallel and antiparallel.

1. Vitally important in 'materials' in biology.
2. Eg silk – silk worms spin the protein fibroin which contains massive regions of antiparallel β pleated sheets.
3. On the other hand collagen – the bone matrix for the mineral formation, tendons – collagen holds everything together.
4. Formed by the triple helix of the tropocollagen protein – 300 nm strands of tropocollagen then align side by side staggered by about 64 nm to form the large structures that exhibit flexibility. The molecules are crosslinked for additional strength and stability. Collagen formation requires vit C (ascorbic acid), lack of this results in break down of tendons – gums bleed, teeth fall out, skin breaks up, blood vessels weaken, death ensues = scurvy** – hence 'Limies'. The use of mustard and water cress***

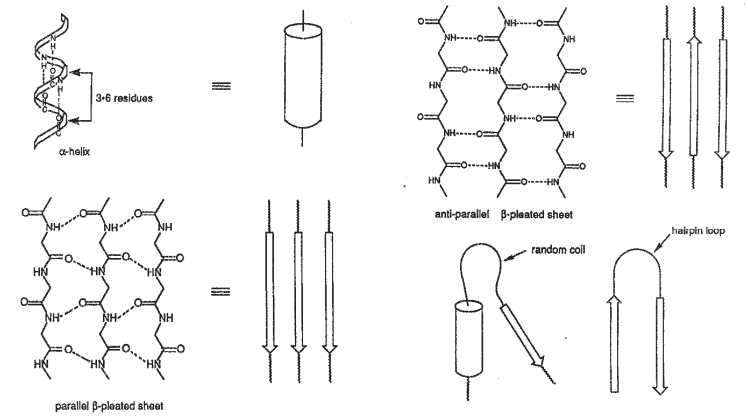
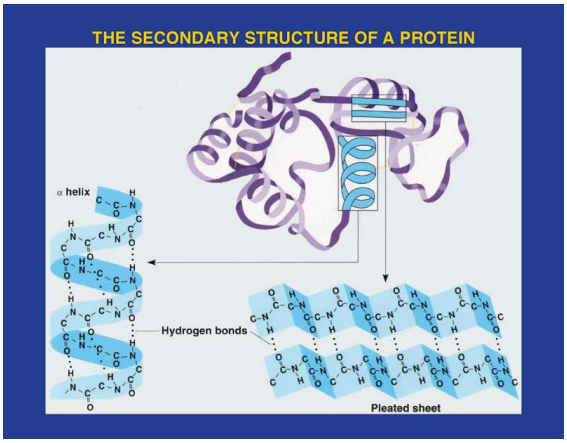


Figure 1.19 Examples of secondary structure and the schematic notation used to represent each structural form. Regions of the peptide backbone that connect α-helices or β-sheets are often called *random coil*. The short sequence linking antiparallel β-sheets is termed a *hairpin loop* or *reverse turn*.

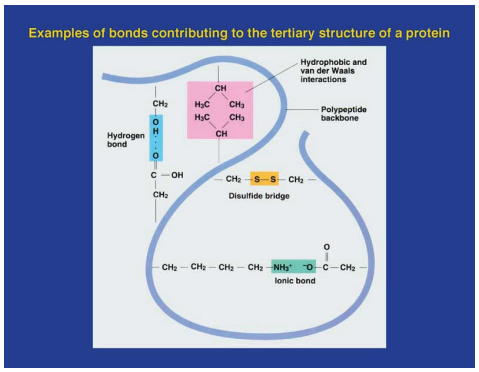
** Scurvy is caused by lack of vitamin C or ascorbic acid. As rich sources of vitamin C , lemon and lime are regarded as foods or exceptional therapeutic value in scurvy. In 1536, the French explorer Jacques Cartier, exploring the St. Lawrence River, used the local natives' knowledge to save his men who were dying of scurvy. He boiled the needles of the Eastern White Cedar to make a tea that was later shown to contain 50 mg of vitamin C per 100 grams. James Lind (1716–1794) first proved scurvy could be treated with citrus fruit in experiments he described in his 1753 book, *A Treatise of the Scurvy*. (From Wikipedia)

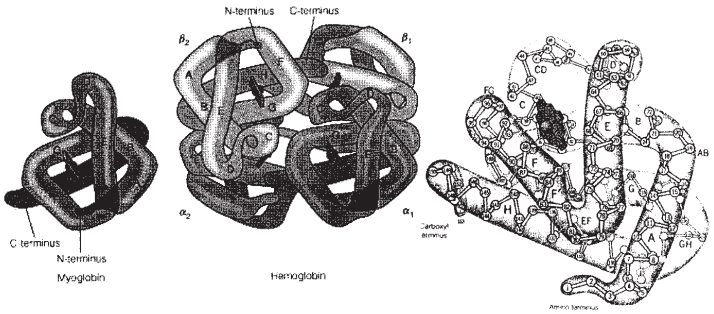
*** http://www.nutrition-and-you.com/watercress.html

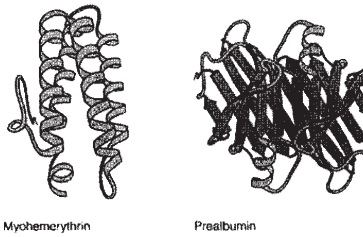


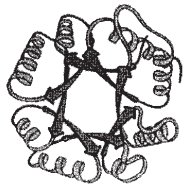
3122

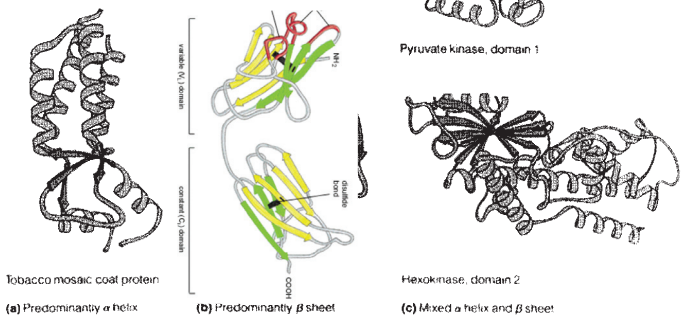
Some examples of structures – identify which structure applies to which protein.











ote – 4 types of bio-interaction – plus 1 (=5)

i) hydrogen bonding,

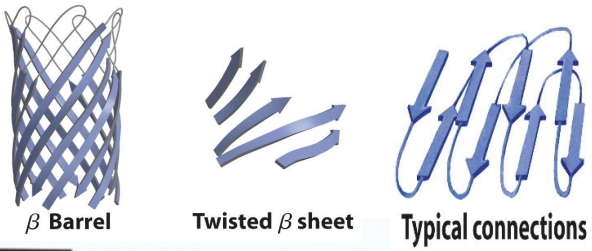
ii) hydrophobic interactions between hydrophobic side chains (always very organic, not ionic – induced dipoles);

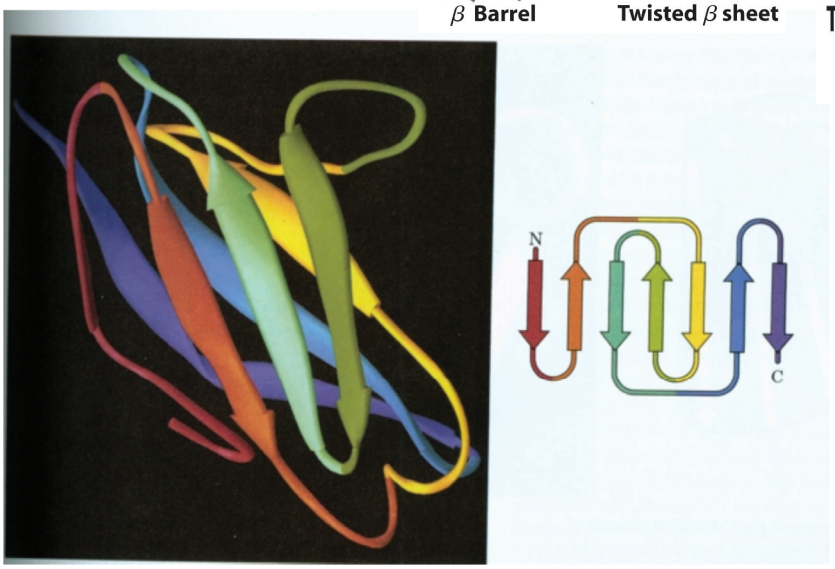
iii) ionic bonds – or electrostatic bonds;

iv) disulfide bridges – the strongest because these are the only true covalent bond that 'cross-links' the protein;

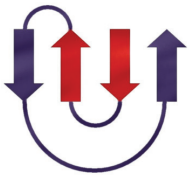
and (v) bonds to metals

Structure is so important for recognition here are some extra views of immunoglobins: Immunological function is based on structural recognition





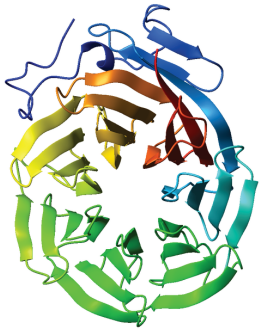
Secondary Structures – examples of β-pleated sheets

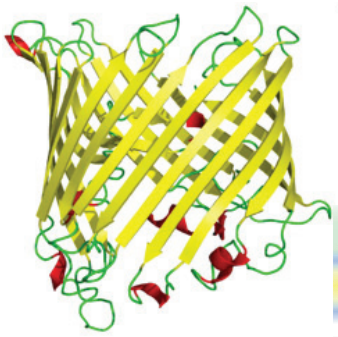


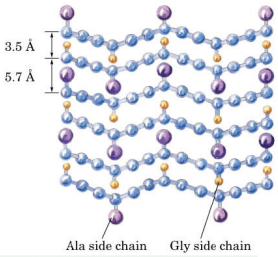
Need a hairpin – very tight turn – use a proline

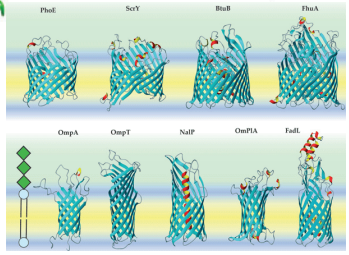
parallel / antiparallel – arrow point towards "C" terminal - - arrows N → C

Examples, silk – best from mulberry trees based on glycine, alanine, serine, with a few glutamic acid, valine, apsartic acid (highly polar side-chanes).









Ribbon diagram of the C-terminal WD40 domain of Tup1 (a transcriptional co-repressor in yeast), which adopts a 7-bladed beta-propeller fold. Ribbon is colored from blue (N-terminus) to red (C-terminus).

EMBO J. 2000 June 15; 19(12): 3016–3027.
Structure of the C-terminal domain of Tup1, a corepressor of transcription in yeast
Elizabeth R. Sprague,[1] Michael J. Redd,[2] Alexander D. Johnson,[2] and Cynthia Wolberger Structure of the C-terminal domain of Tup1, a corepressor of transcription in yeast. Elizabeth R. Sprague,Michael J. Redd, Alexander D. Johnson, and Cynthia Wolberger
http://www.molecularstructure.org/entry.php?pdb=1erj

A canonical beta-barrel protein, a sucrose-specific porin from the bacterium Salmonella typhimurium, viewed from the side. Porins are transmembrane proteins with hollow centers through which small molecules can diffuse.

http://www.youtube.com/watch?v=WuFBTQs8EWM&feature=player_embedded

EG- Schulz GE: The structure of bacterial outer membrane proteins. *Biochim Biophys Acta* 2002, 1565(2):308-317.

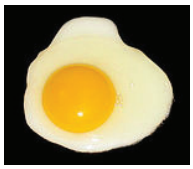
All structure can be lost – called denaturation.

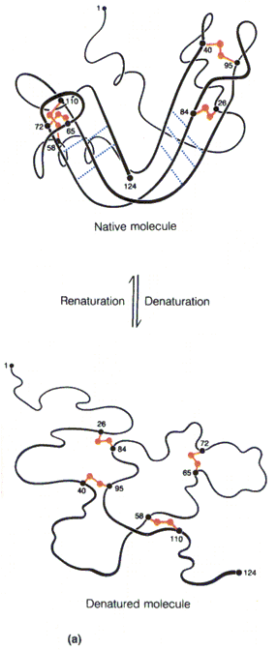
We can identify the change in structure from a 'melting curve' if we have a technique that we can use to measure the structural change as a function of an applied 'insult'.

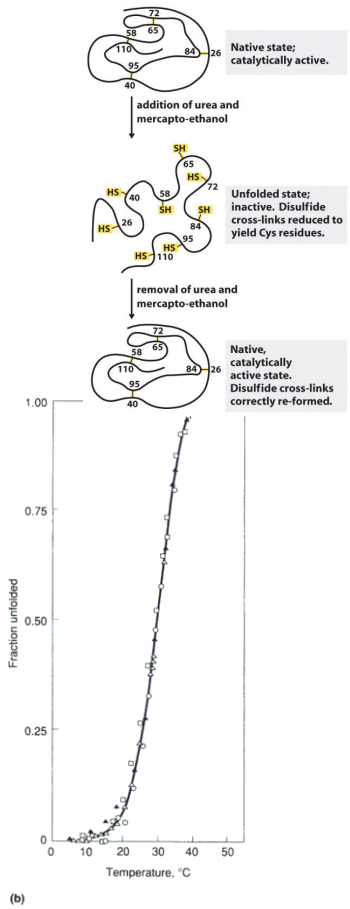
Heating a solution is often sufficient.

Urea, mercaptoethanol (stinky) or guanidinium chloride are reversible denaturants – increase the concentration of these agents in solution and the protein denatures – decrease the concentration and thr native form returns.

If the cys-S-S-cys bonds are broken, then this results in irreversible denaturation – sometimes the tangle can reform its original structure – reversible denaturation.







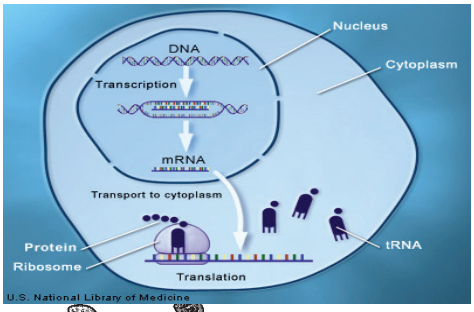
3127
3095

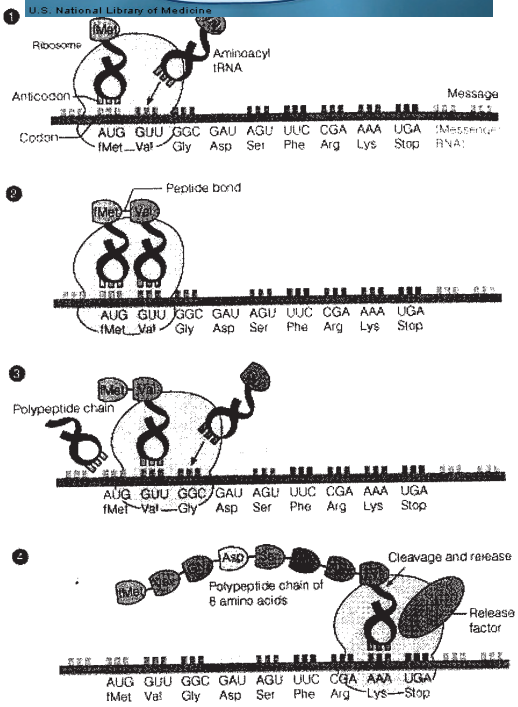
See more detail including information about the tRNA, mRNA and DNA role on slide 24

How do peptide chains form in vivo?

1. ΔG = about +10kJ/mol for formation – that is the break up of the peptide bond is favoured, hydrolysis.
2. This is the thermodynamic view, the kinetic view shows this to be a very slow reaction.
3. So, we will all break up in time. Faster in boiling HCl or with proteolytic enzymes present. (In our digestive tracts for example, trypsin, chymotrypsin, thrombin, carboxypeptidase A, and others – think of pineapple, mmmm – dissolved lips from the bromelain).
4. Luckily, enzymes glue the amino acids back together, but how in the first place does the peptide chain form?
5. Triplets of nucleotides (called codons) that span the DNA sequence, code mRNA, which is then the template as shown. This way the 4 nucleotides can code for 64 different combinations of amnio acid (4^3), but only 20 are used, so that leaves redundancy.
6. AUG is the start signal – the N- terminus aa = N-formylmethionine
7. UAA, UAG, UGA code the stop, for the C-terminus aa.
8. The 3-letter code on DNA is transferred to mRNA. There are 3 steps in the ribosome – initiation, elongation, release.
9. In the ribosome (this massive molecule comprises several subunits and involves a large number of protein 'cofactors') tRNA with a specific aa on one end and the anticodon complementary code on the other are brought up to the mRNA. The ribosome moves along the mRNA strand as shown. The peptide bond forms, the old tRNA drops off and the process is repeated. The energy source of ATP -> ADP + P_i -> AMP + P_i

See L&B p 50-55 (rather more extensive than here but interesting about cloning.)





Classes of proteins – proteins with specific structures

All alpha helix - myoglobin – tropomysoin – many membrane bound proteins -

All beta-pleated sheet – prealbumin, concanavalin A

Mixed alpha -beta regions –

Membrane proteins – transporters

Irregular or Others – proteins without a specific structure – metallothionein.

Classes of proteins – – proteins that carry out specific functions

Structural - fibrous **collagen** – keratin

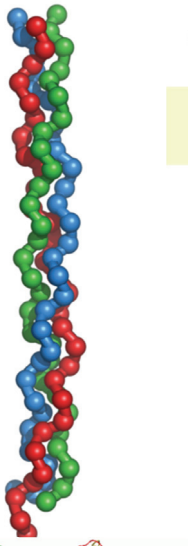
Motor proteins – actin - – myosin – muscle contraction

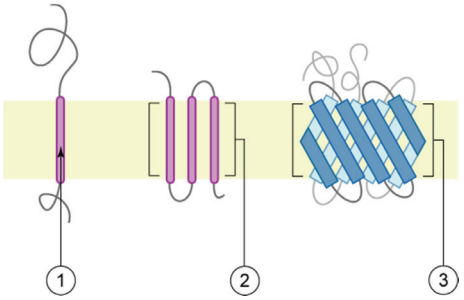
Transport – oxygen in hemoglobin

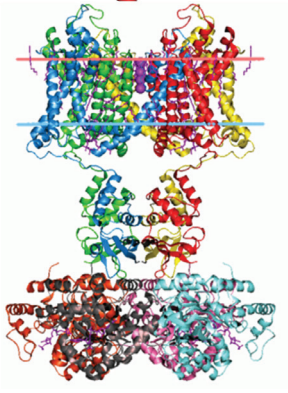
Storage – iron in ferritin, metallothionein

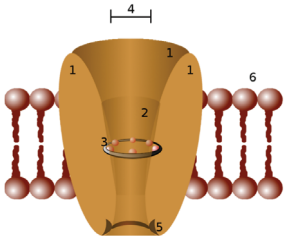
Membrane – ion channels – add permeability to a membrane for small molecules

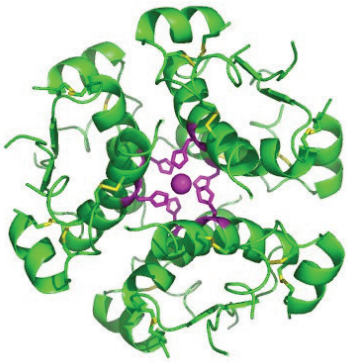
Signalling proteins – insulin – in the quaternary hexameric structure the metal is??







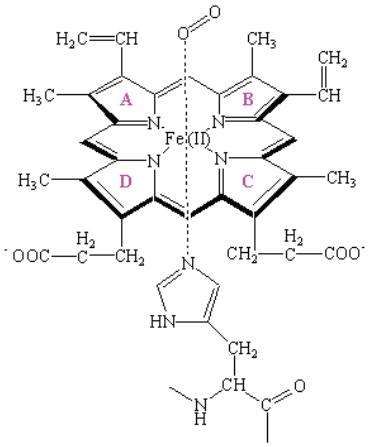




Crystal structure of Potassium channel KvAP. Calculated hydrocarbon boundaries of the lipid bilayerare indicated by red and blue dots

Metallproteins & Metalloenzymes

Common metals: Fe^{2+}, Fe^{3+}, Cu^{1+}, Cu^{2+}, Zn^{2+}, Mg^{2+}, Mn^{2+**}, Co^{3+}, Mo^{5+**}



Iron-containing proteins -

Hemoglobin and mitochondrial electron transport - to carry O_2.

Cytochromes - Heme proteins- one electron transfer agents. Fe2+ - Fe3+

Iron-sulfur proteins - One electron transfer agents. Fe2+ - Fe3+

Copper-containing proteins - One electron transfer agents. Hemocyanin***

Zinc-containing proteins - e.g. carbonic anhydrase - Hydration of CO_2. Three His residues coordinate Zn^{2+}.

Manganese-containing proteins - e.g. arginase

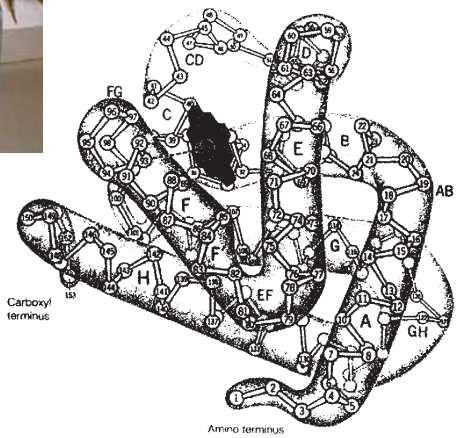
Molybdenum-containing proteins - nitrate reductase

**for completeness- -not on tests

L&B p 286 for a different image

Fe-PPIX is the heme in this heme protein

***blue-pigmented oxygen carriers, in arthropods and molluscs.

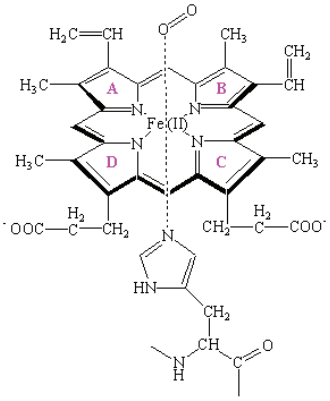


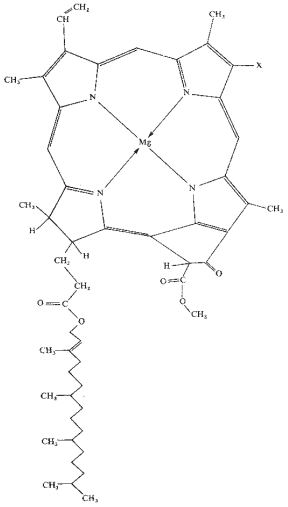
Cofactors

To remind ourselves of some of these important molecules

Porphyrins – many slightly different ones are found in proteins:

Protoporphyrin IX in myoglobin and hemoglobin (as well as a number of other enzymes, HRP, catalase) is called heme b.

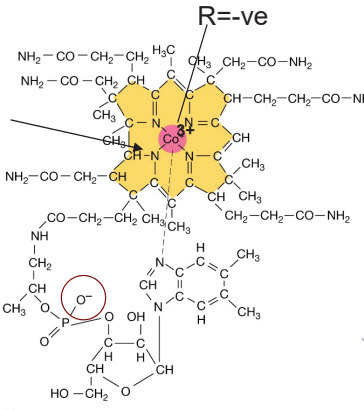


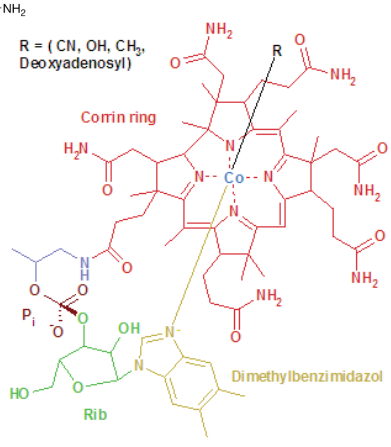


Vitamin (cofactor) B group*:

Corrin in Vit B12 -uses R = CN^- on the Co^{3+} - see below for other R groups*. Methylcobalamin is a cofactor in methionine synthase (MTR) – the enzyme that connects both methionine synthesis and folic acid synthesis. (Lack → spina bifida – disease involving myelin coating of nerves)

Chlorin→ chlorophyll – not really an enzyme – the chlorophyll acts as a photon trap – ejecting an electron.





*cobalt-containing compounds known as corrinoids. The principal cobalamins are cyanocobalamin, hydroxocobalamin and the two coenzyme forms of vitamin B_{12}, methylcobalamin and 5-deoxyadenosylcobalamin (adenosylcobalamin).

1150

Four of the six coordination sites of the Co(III) are provided by the corrin ring, and a fifth by a dimethylbenzimidazole group. The sixth coordination site, the center of reactivity, is variable, in tablet for it is -CN, in vivo, a hydroxyl group (-OH), a methyl group (-CH_3) or a 5'-deoxyadenosyl group – see the B12 unit coming.

Ref for pictures: http://en.wikipedia.org/wiki/B12

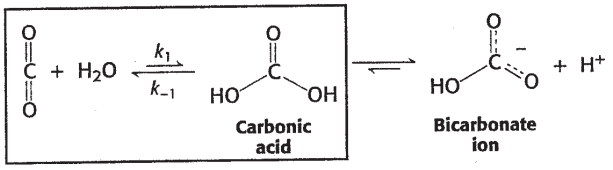
(a) dimethylbenzimidazole

Note the sugar phosphate chain (ribose)

Carbonic Anhydrase

3126

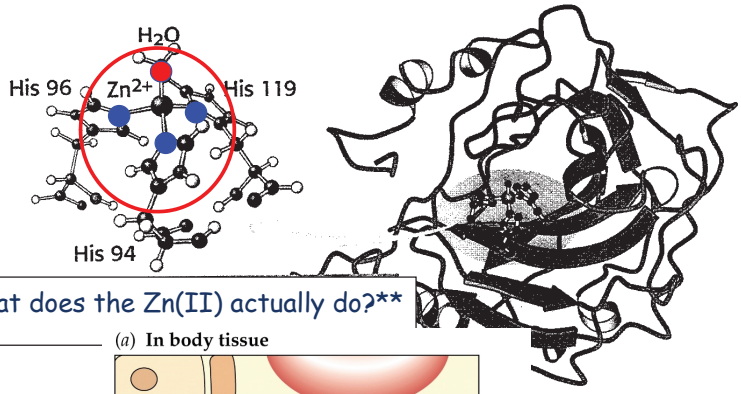
Cofactors like Zn also activate proteins and enzyn
fact, many other biologically important molecules,
ADP (see below), use just a single metal.



k_1 = .0027 $M^{-1}s^{-1}$, k_{-1} = 50 s^{-1}, and K_{eq} = 5 x 10^{-5}

Carbonic anhydrase was the first enzyme to
discovered that contained Zn(II). Why so h

Note the chemistry that takes place (L&B p 270) –
the Zinc unit coming.

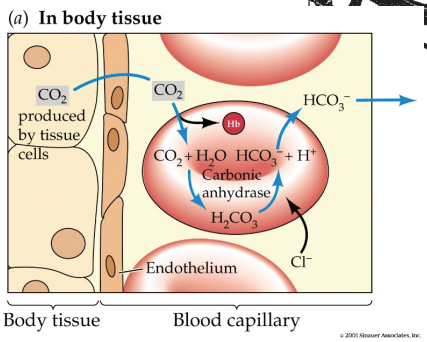


The Zn(II) is bound by 3 HIS and 1 water

Does this fit our model?

L&B p 267-271; read on about alcohol dehydrogenas
because we will come back to these 2 when discuss
Zn(II) later.

What does the Zn(II) actually do?**

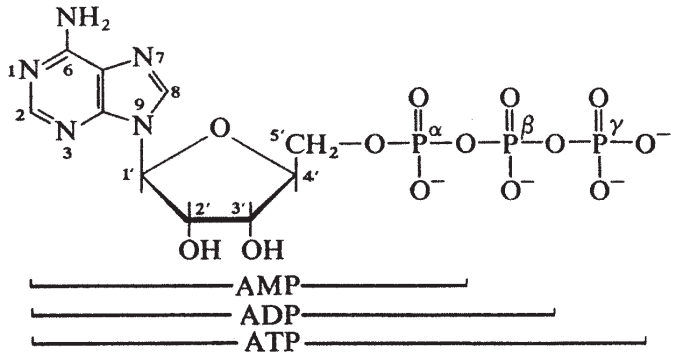


*d^{10} – colourless- and Zn(II) is everywhere

** The Zn(II) 'activates' the water – makes it OH^- - this attacks the CO_2 forming HCO_2^-

3075a

The source of energy is the breaking
of the O-P-O bond



The net energy released upon ATP
hydrolysis to ADP + P_i?

$[ATP]^{4-} + H_2O \rightarrow [ADP]^{3-} + HPO_4^{2-} + H^+$

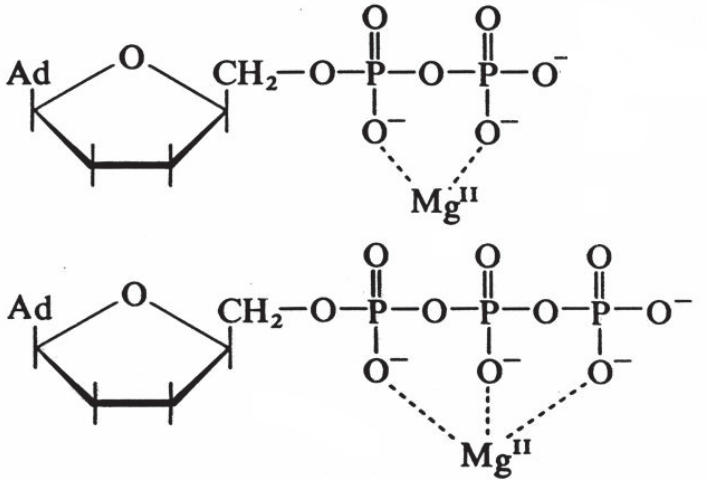
or:

$MgATP^{2-} + H_2O \rightarrow MgADP^- + P_i + H^+$

-7.3 kcal/mol (1M conc) or -30.5 kJ/mol
(-12.9 kcal/mol (at 1 mM conc.))

(to convert cal to J, multiply by 4.18)

(Ad = adenine group attached to the ribose sugar)



3075b

Formation of the polymeric nucleotide chain

A sugar phosphate chain with the nucleotide bases always in the same position.

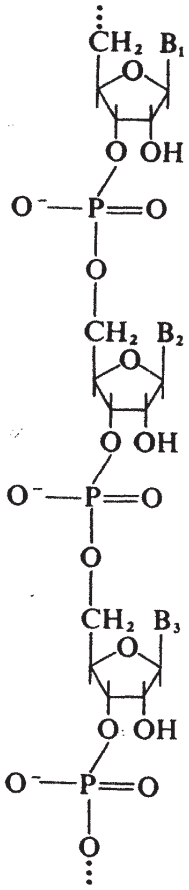
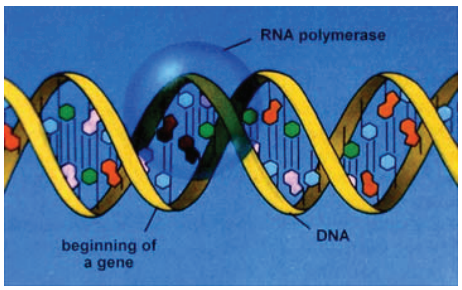


FIGURE 2.2. Sequence of a polynucleotide.

But what was the structure?

The chain was known, the fact that the nucleotides existed in exactly the same ratios was known but what did this mean?



MOLECULAR STRUCTURE OF NUCLEIC ACIDS

A Structure for Deoxyribose Nucleic Acid

WE wish to suggest a structure for the salt of deoxyribose nucleic acid (D.N.A.). This structure has novel features which are of considerable biological interest.

A structure for nucleic acid has already been proposed by Pauling and Corey[1]. They kindly made their manuscript available to us in advance of publication. Their model consists of three intertwined chains, with the phosphates near the fibre axis, and the bases on the outside. In our opinion, this structure is unsatisfactory for two reasons: (1) We believe that the material which gives the X-ray diagrams is the salt, not the free acid. Without the acidic hydrogen atoms it is not clear what forces would hold the structure together, especially as the negatively charged phosphates near the axis will repel each other. (2) Some of the van der Waals distances appear to be too small.

Another three-chain structure has also been suggested by Fraser (in the press). In his model the phosphates are on the outside and the bases on the inside, linked together by hydrogen bonds. This structure as described is rather ill-defined, and for this reason we shall not comment on it.

We wish to put forward a radically different structure for the salt of deoxyribose nucleic acid. This structure has two helical chains each coiled round the same axis (see diagram). We have made the usual chemical assumptions, namely, that each chain consists of phosphate di-ester groups joining β-D-deoxyribofuranose residues with 3′,5′ linkages. The two chains (but not their bases) are related by a dyad perpendicular to the fibre axis. Both chains follow right-handed helices, but owing to the dyad the sequences of the atoms in the two chains run in opposite directions. Each chain loosely resembles Furberg's[2] model No. 1; that is, the bases are on the inside of the helix and the phosphates on the outside. The configuration of the sugar and the atoms near it is close to Furberg's 'standard configuration', the sugar being roughly perpendicular to the attached base. There is a residue on each chain every 3·4 A. in the z-direction. We have assumed an angle of 36° between adjacent residues in the same chain, so that the structure repeats after 10 residues on each chain, that is, after 34 A. The distance of a phosphorus atom from the fibre axis is 10 A. As the phosphates are on the outside, cations have easy access to them.

The structure is an open one, and its water content is rather high. At lower water contents we would expect the bases to tilt so that the structure could become more compact.

The novel feature of the structure is the manner in which the two chains are held together by the purine and pyrimidine bases. The planes of the bases are perpendicular to the fibre axis. They are joined together in pairs, a single base from one chain being hydrogen-bonded to a single base from the other chain, so that the two lie side by side with identical z-co-ordinates. One of the pair must be a purine and the other a pyrimidine for bonding to occur. The hydrogen bonds are made as follows: purine position 1 to pyrimidine position 1; purine position 6 to pyrimidine position 6.

If it is assumed that the bases only occur in the structure in the most plausible tautomeric forms (that is, with the keto rather than the enol configurations) it is found that only specific pairs of bases can bond together. These pairs are: adenine (purine) with thymine (pyrimidine), and guanine (purine) with cytosine (pyrimidine).

In other words, if an adenine forms one member of a pair, on either chain, then on these assumptions the other member must be thymine; similarly for guanine and cytosine. The sequence of bases on a single chain does not appear to be restricted in any way. However, if only specific pairs of bases can be formed, it follows that if the sequence of bases on one chain is given, then the sequence on the other chain is automatically determined.

It has been found experimentally[3,4] that the ratio of the amounts of adenine to thymine, and the ratio of guanine to cytosine, are always very close to unity for deoxyribose nucleic acid.

It is probably impossible to build this structure with a ribose sugar in place of the deoxyribose, as the extra oxygen atom would make too close a van der Waals contact.

The previously published X-ray data[5,6] on deoxyribose nucleic acid are insufficient for a rigorous test of our structure. So far as we can tell, it is roughly compatible with the experimental data, but it must be regarded as unproved until it has been checked against more exact results. Some of these are given in the following communications. We were not aware of the details of the results presented there when we devised our structure, which rests mainly though not entirely on published experimental data and stereochemical arguments.

It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material.

Full details of the structure, including the conditions assumed in building it, together with a set of co-ordinates for the atoms, will be published elsewhere.

We are much indebted to Dr. Jerry Donohue for constant advice and criticism, especially on interatomic distances. We have also been stimulated by a knowledge of the general nature of the unpublished experimental results and ideas of Dr. M. H. F. Wilkins, Dr. R. E. Franklin and their co-workers at King's College, London. One of us (J. D. W.) has been aided by a fellowship from the National Foundation for Infantile Paralysis.

J. D. WATSON
F. H. C. CRICK
Medical Research Council Unit for the Study of the Molecular Structure of Biological Systems,
Cavendish Laboratory, Cambridge.
April 2.

[1] Pauling, L., and Corey, R. B., *Nature*, 171, 346 (1953); *Proc. U.S. Nat. Acad. Sci.*, 39, 84 (1953).
[2] Furberg, S., *Acta Chem. Scand.*, 6, 634 (1952).
[3] Chargaff, E., for references see Zamenhof, S., Brawerman, G., and Chargaff, E., *Biochim. et Biophys. Acta*, 9, 402 (1952).
[4] Wyatt, G. R., *J. Gen. Physiol.*, 36, 201 (1952).
[5] Astbury, W. T., Symp. Soc. Exp. Biol. 1, Nucleic Acid, 66 (Camb. Univ. Press, 1947).
[6] Wilkins, M. H. F., and Randall, J. T., *Biochim. et Biophys. Acta*, 10, 192 (1953).

This figure is purely diagrammatic. The two ribbons symbolize the two phosphate–sugar chains, and the horizontal rods the pairs of bases holding the chains together. The vertical line marks the fibre axis

3076b

The critically important discovery was that that the special pairs formed and only these pairs

TA and GC

L&B 56-64 R-M 39-45

Approximate energy of a hydrogen bond? 2 - 10 kcal/mol depending on the local environment:

2-3 kcal/mol (8-12 kJ/mol) in water

But! 6-7 kcal/mol (29 kJ/mol) in membranes

Compare this with the energy of a covalent bond?

50 - 90 kcal/mole (250-400 kJ/mol)

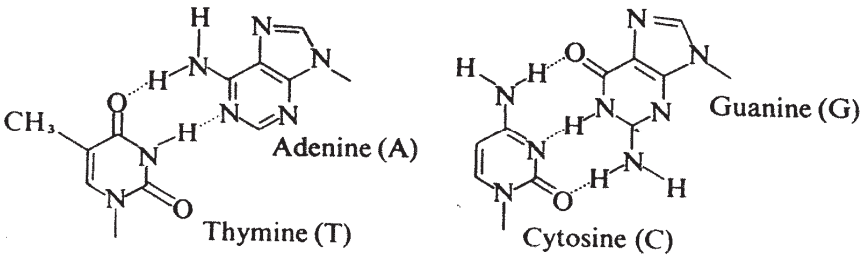
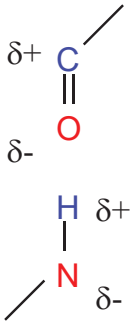
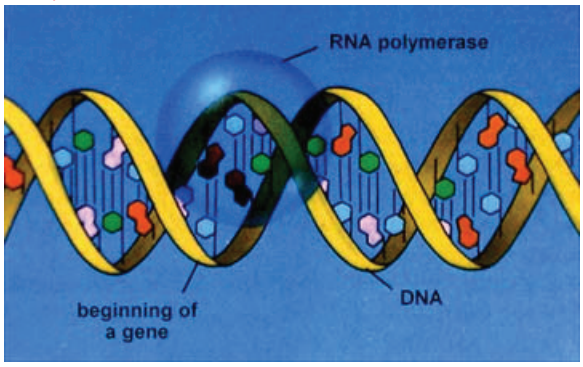


FIGURE 2.3. Pair formation between purine and pyrimidine bases in DNA.



1071a

The double helix then forms with the chains running antiparallel



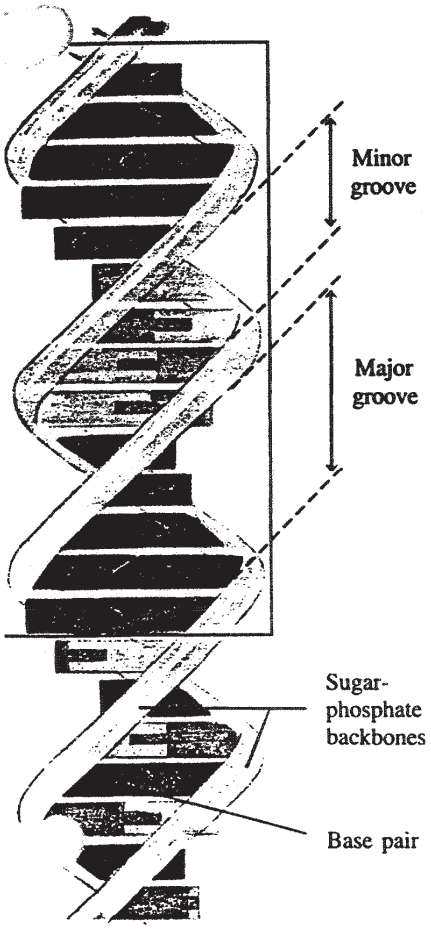
3.6 nucleotides per turn.

5.4 Å rise in one turn of a helix.

And the energy net energy released upon ATP hydrolysis to ADP + Pi?

-7.3 kcal/mol (1M conc) or 30.5 kJ/mol

(-12.9 kcal/mol (1 mM conc.))



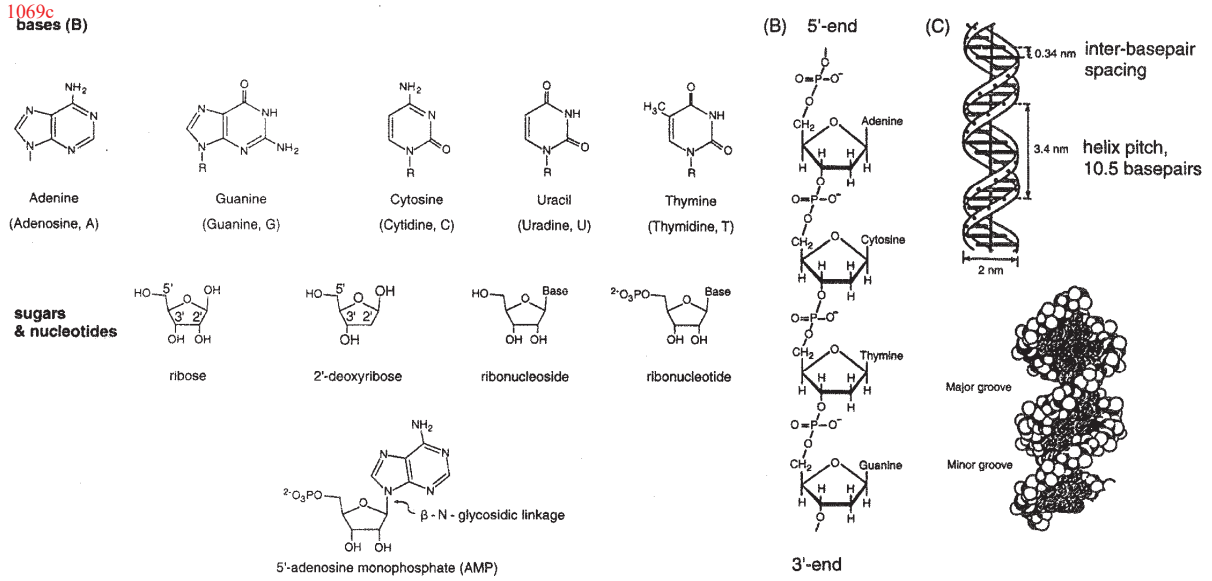
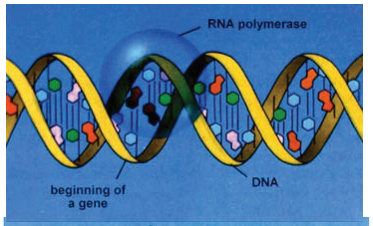
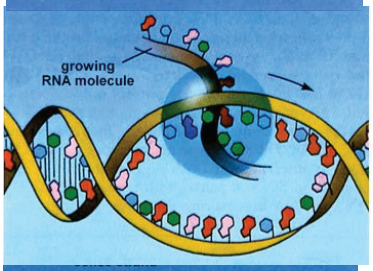
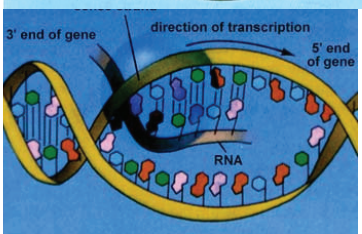
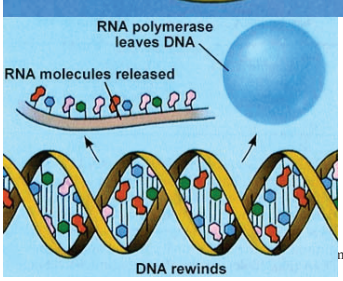


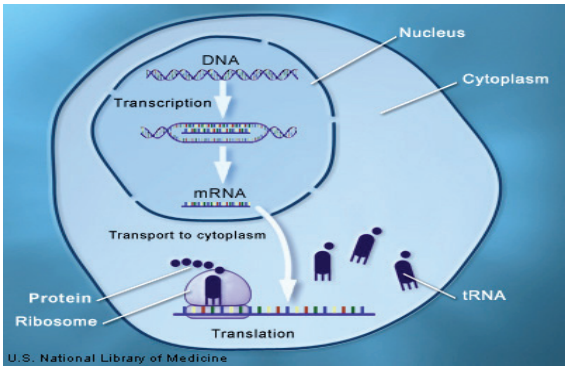
Figure 1.20 Structural units of the nucleic acids. (A) Purine and pyrimidine bases, ribose sugar, and nucleotides. The names in parentheses below the bases are the names of the corresponding nucleosides or nucleotides (i.e., base + ribose-phosphate). The 2′, 3′, and 5′ carbons on ribose are indicated. (B) Base attached to the deoxyribose–phosphate backbone. A single strand of DNA. Double-stranded DNA is formed by specific hydrogen bond patterns formed between complementary base pairs (see Figure 1.21). (C) Structural features of B-DNA. The vertical rise of 3.4 nm for a complete 360° turn is accommodated within 10.5 base units. (Adapted from R. L. P. Adams, J..T. Knowler, and D. P. Leader, *The Biochemistry of the Nucleic Acids*, Chapman and Hall, 1986).

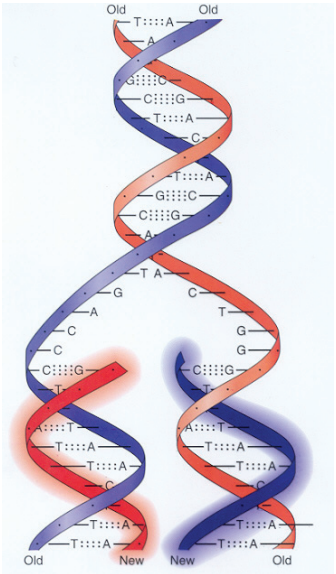








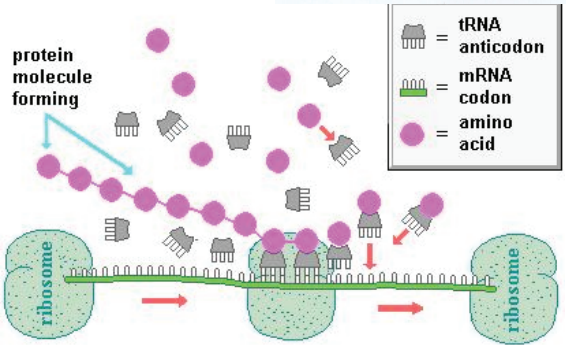




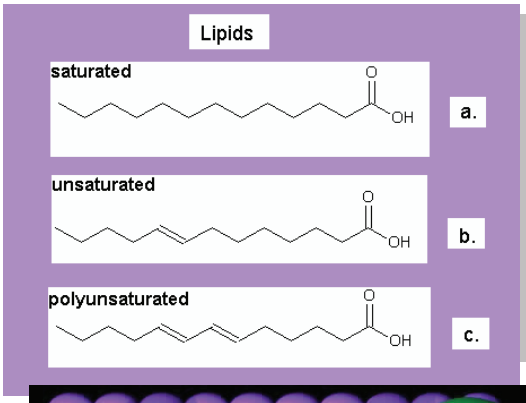
Synthesis of proteins – for a second time but with more detail

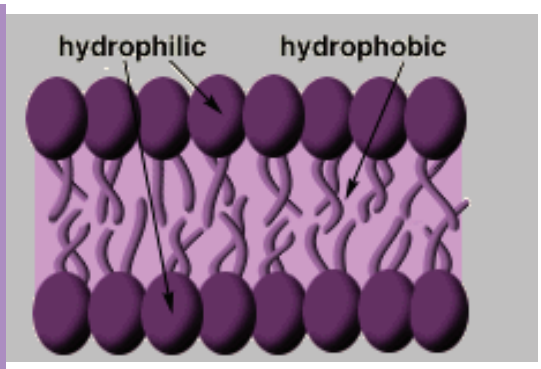
1. Replication of the genetic code held on DNA
2. Formation of m-RNA
3. Use of T-RNA in the to code for amino acids (3 bases at a time needed)

All in the ribosome

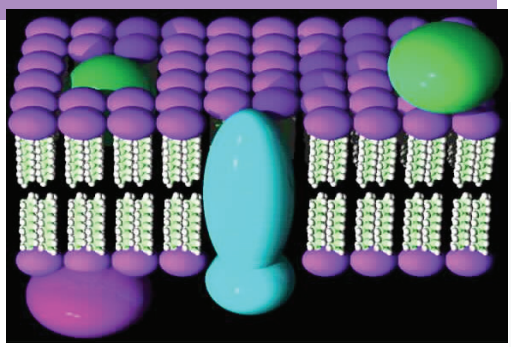


Need to memorize differences – be able to recognise each of these 3.



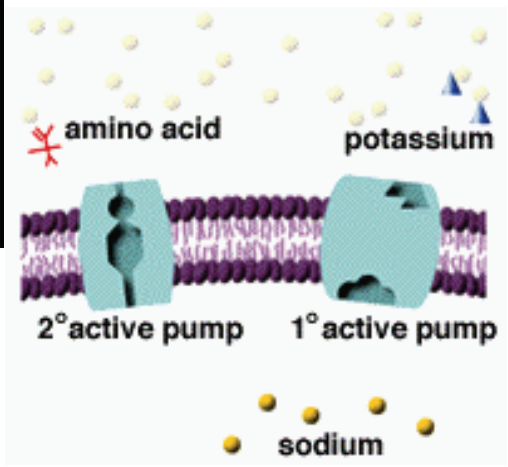


Saturated fats raise total blood cholesterol as well as LDL cholesterol (the bad cholesterol). Saturated fats are mainly found in animal products such as meat, dairy, eggs and seafood. Some plant foods are also high in saturated fats such as coconut oil, palm oil and palm kernel oil.



Monounsaturated fats lower total cholesterol and LDL cholesterol (the bad cholesterol) and increase the HDL cholesterol (the good cholesterol). Nut, canola and olive oils are high in monounsaturated fats.

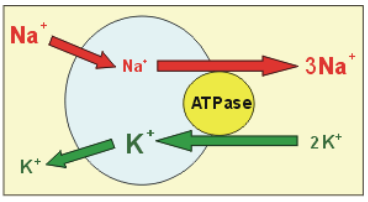
Polyunsaturated fats also lower total cholesterol and LDL cholesterol. Seafood like salmon and fish oil, as well as corn, soy, safflower and sunflower oils are high in polyunsaturated fats. Omega 3 fatty acids belong to this group.

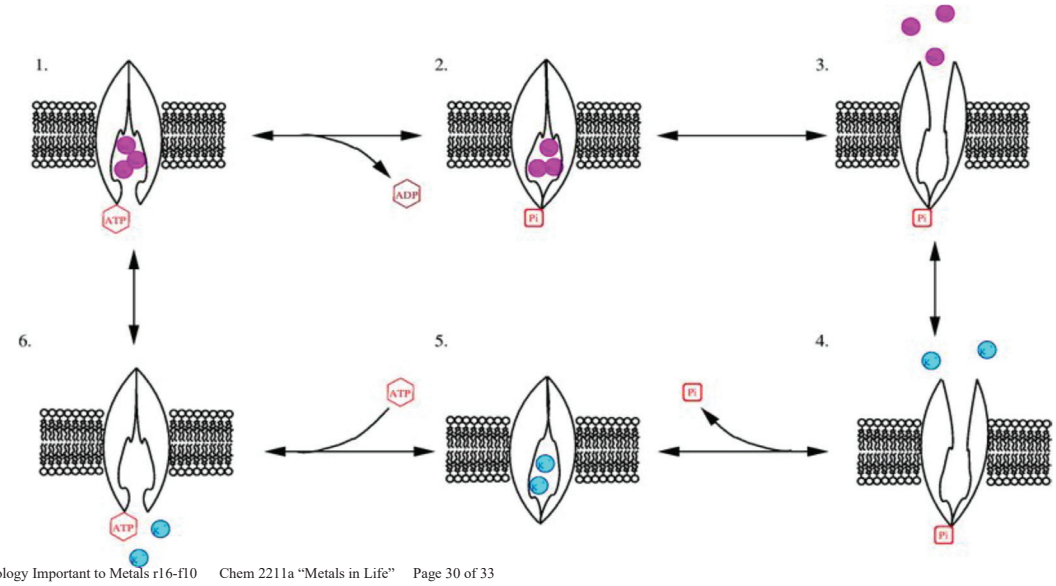


So, how does the simplest of pumps work?

Pumping Na^+ out and K^+ in to the cell.

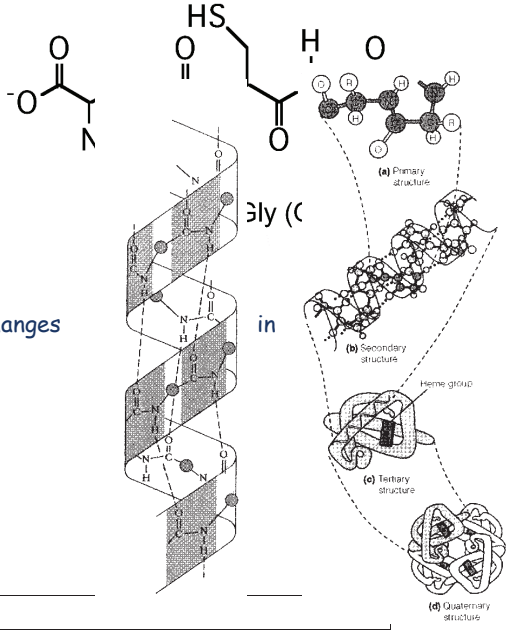
Note the consumption of energy – fuel - $2K^+$ pumped, $3Na^+$ pumped out of the cell. Outside is high in Na^+ but low in K^+.





Key points from this unit	
1	Must know the overall structure of the 20 essential amino acids - - and the detailed structures of all metal binding amino acids – know how many are required in the diet (10), how many mammals can synthesize (10) Know the classes of amino acid – and be able to recognise from each class
2	Peptide chain formation – know the overall steps Know the 4 structural terms and have examples of proteins with at least 3y and 4y structures dominating Know the 4 types of interaction Know about denaturation – breaking S-S bonds
3	Know the common metals for metalloproteins – and know what enzymes use them and for what purpose Check how carbonic anhydrase works
4	ATP-ADP-AMP know the energies involved – draw ATP Draw the sugar phosphate backbone Know the hydrogen bonded base pairs – be able to draw the pairs Know the basics of cellular chemistry – where does DNA reside? Where does protein synthesis take place? Know how cell membranes are constructed

Glutathione

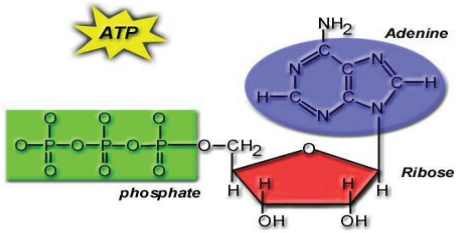


Study questions from the lectures to date and from th	
The 4 structures in protein biology An example of the progression from primary to secondary to te quaternary structure for hemoglobin Notice how the hydrogen bonds hold the α helix together. Denaturing agents break these bonds resulting in spectacular changes properties, eg egg white after boiling. 3004	
Lectures	
S-L	Ch 3 - p 43-50; 56-59; 69-72; Qu 1.; Note the Desferrioxamine-B structure – no need to memorize except know that is contains many C=O and N-OH groups. Hard or soft? Which metal does it bind?
R-M	
K-S	No specific pages

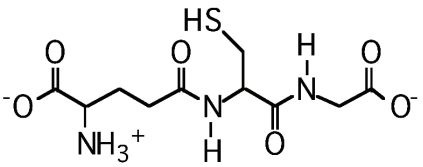
3095

Some important bio-molecules and processes to really know:
(Include the 4 nucleotide bases from above – no need to be able to draw the bases - but you need to recognise – identify each.)

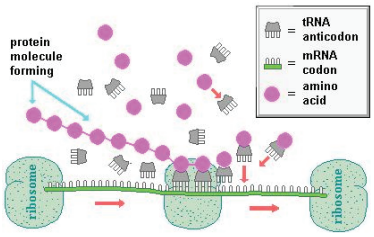
And the metal binding amino acids: His – delta N or epsilon N. (δ or ε) - Asp – with the possibilities we saw earlier for different bonding modes through the $–COO^-$ group. –Cys - Met - Glu (note 1 CH_2 extra in Glu c.f. Asp) - Tyr



Glutathione



γ-Glu-Cys-Gly (GSH)



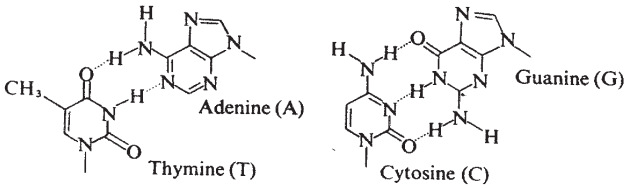


FIGURE 2.3. Pair formation between purine and pyrimidine bases in DNA.