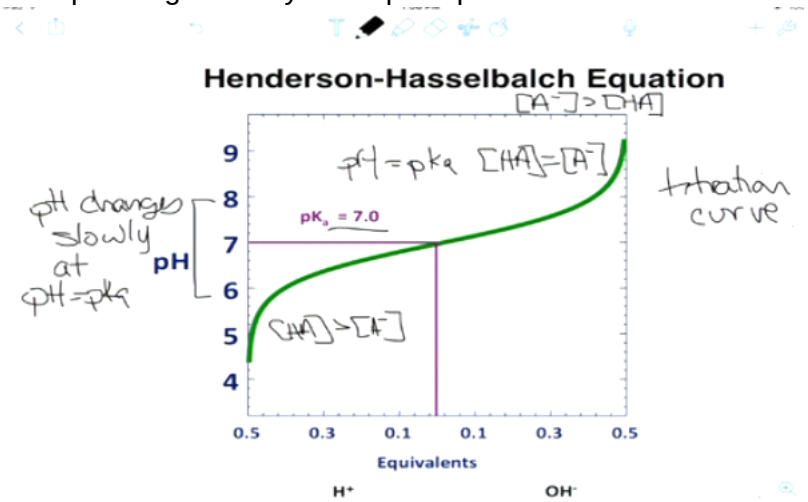
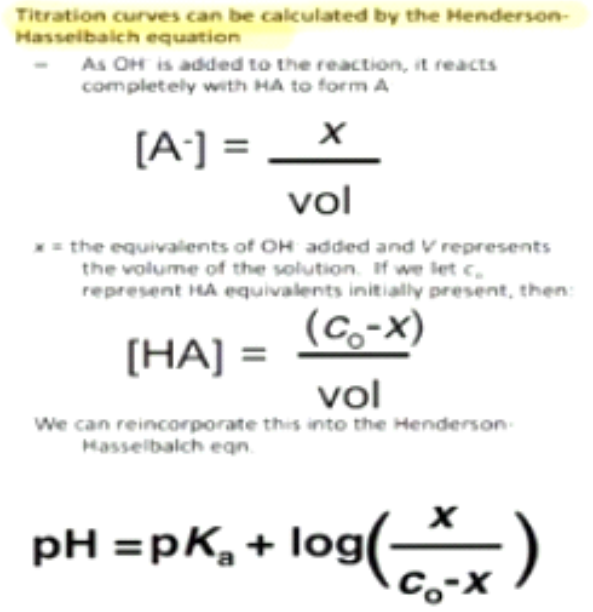
Amino Acids - Lecture 3 and 4

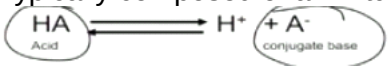
January 17, 2017 1:01 PM

- Henderson - Hasselbach
  - Relates pH to the pkA to describe the ratio to of the acid to the conj. Base
- At the pH to the pkA point - relatively low change in pH with adding assdfcid
  - Outside of this range, acid or base causes a rapid change
  - pH canges slowly when pH = pKa
- 

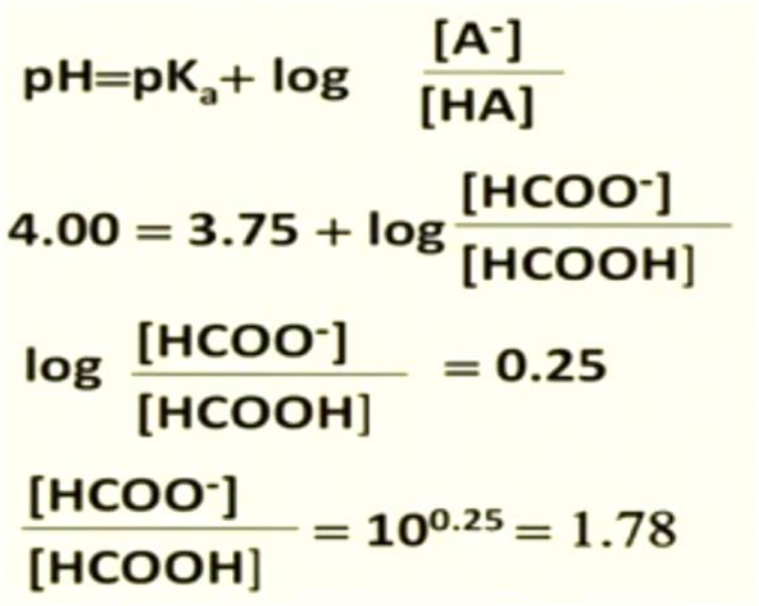
- pKa can be determined by the midpoint of the titration curve
- Titration curves are determined using the approach shown
- 

- Think of it as adding in OH - reacts with HA to produce A-
- If C_o = initial concentration of HA

Buffers

- Resist changes in their pH as acid [H+] or base {OH-] is added, essential for biochemical system as they may undergo protonation if pH is not controlleddsf
- Typically composed of a mixture of weak acids and conj. Weak base
- 

- Within one pH unit of one pKa that is the range that is best at (pKa of 7, buffer range of 6 to 8)
- Therefore, can determine which bases and acids to use based on the required buffer
- Required: 0.1M buffer at pH 4.00

- 

  - This just tell us the ratio of acid to conj. Base needed for a pH of 4.00
  - 1.0M buffer is needed - total concentration of the acid and the base together (the species in equilbrium are formice acid and formate)
    - Let [HCOO-] = a and [HCOOH] = b
    - a/b = 1.78 and a+b =0.1 by substitution
    - a = 0.064 M and b = 0.036 M
  - To make the buffer - create a more concentrated solution of the acid and conj. Base - at 1M solution of each, and then dilute down the mixture to the required molarity of the solution
    - 64mL of 1M formic acid and 36 mL of 1M sodium formate and dilute to 1L with water

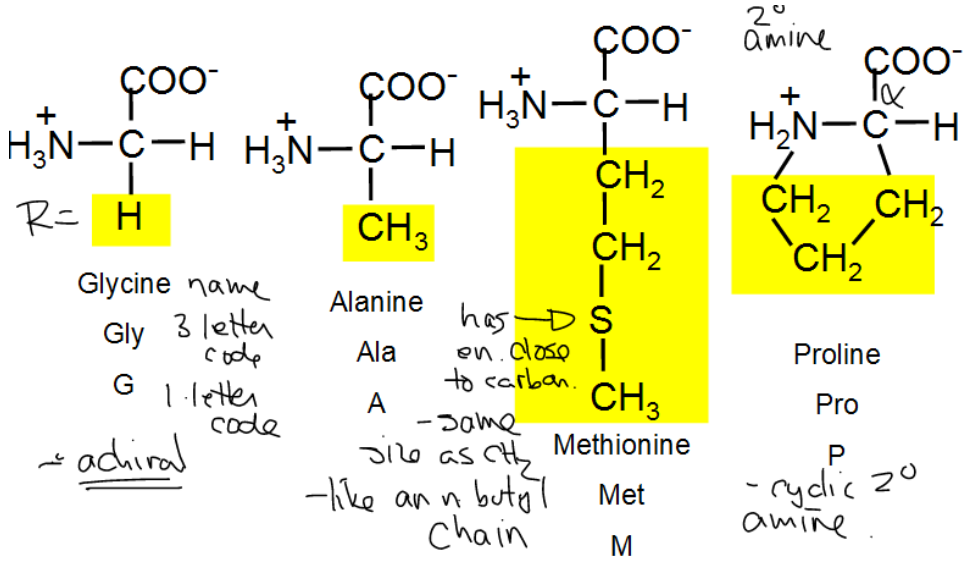
Amino Acid

- 20 standard amino acids
- 3 main groups based on side chains
  - Nonpolar (hydrophobic side chain)
    - Simple aliphatic, aromatics, proline
  - Uncharged polar side chain
    - Alcohol, thiols, amides
  - Charged polar side chains
    - Acidic, basic
- Different non-covalent forces
- Some overlap between groups - depending on the context it is in
- Limited number of functional groups
- Most amino acids are chiral (non-superimposable, mirror images)
  - All except glycine are chiral
    - L-Alanine is a building block, D-alanine is never found in proteins but found in some structures
  - The central Carbon is sp3 hybridized - 4 different groups bonded to it, said to be stereogenic - stereogenicity is a property of atoms, so it is more correct ot describe the carbon as a stereogenic carbon rather than a chiral carbon (Chirality is a molecular property not an atomic propoerty)
  - D-Alanine is found in the peptidoglycan of bacterial cell walls
- Enantiomers
  - Stereogenic centres with the descriptors of R,S (Cahn-Ingold-Prelog convention) - R - turns to the right
  - Fisher-Rosanoff convention is used (use D and L) same as the R and S but is slightly different
    - All proteogenic amino acids are L, in CIP some are S and some R
    - CIP assigns priority based on the atomic number
    - All proteogenic amino acids are S except for cysteine
  - Fisher-Rosanoff - things must be drawn in the Fisher projection
    - Highest oxidation state carbon put at the top - always drawn at ninety degree bonds
    - A pair out of the vertical bonds going into the page, a pair of horizontal bonds

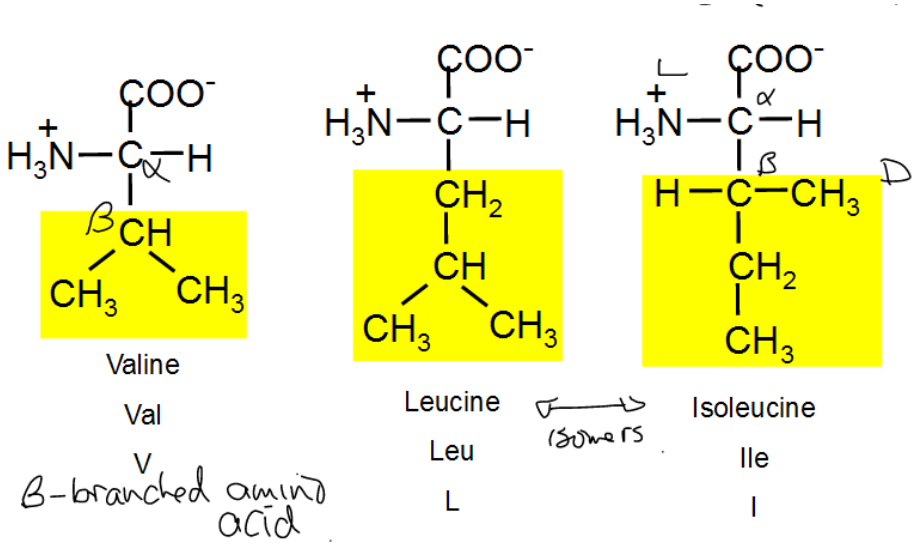
going out of the page
        - Draw the longest linear chain of carbons going down
        - Substituent then put on the horizontal - if the highest priority substituent is on the left side - L if the highest priority substituent is on the right side - D
    - Summary: 1) draw highest oxidation state carbon at the top, 2) extend the chain down, 3) on the horizontal - highest priority substituent on left - L, if on right - D

Aliphatic non-polar

- Gly, Ala, Met, Val, Pro, Leu, and Ile
- 3 different names, full name, 3 letter code, and the single code
- Methionine
    - Contains sulfur - which is a heteroatom but it is considered non-polar
    - Sulfur has an EN similar to carbon so the S -C bonds are not polar bonds
    - Sulfur is also very big - behaves as a CH2 group - like an n-butyl chain

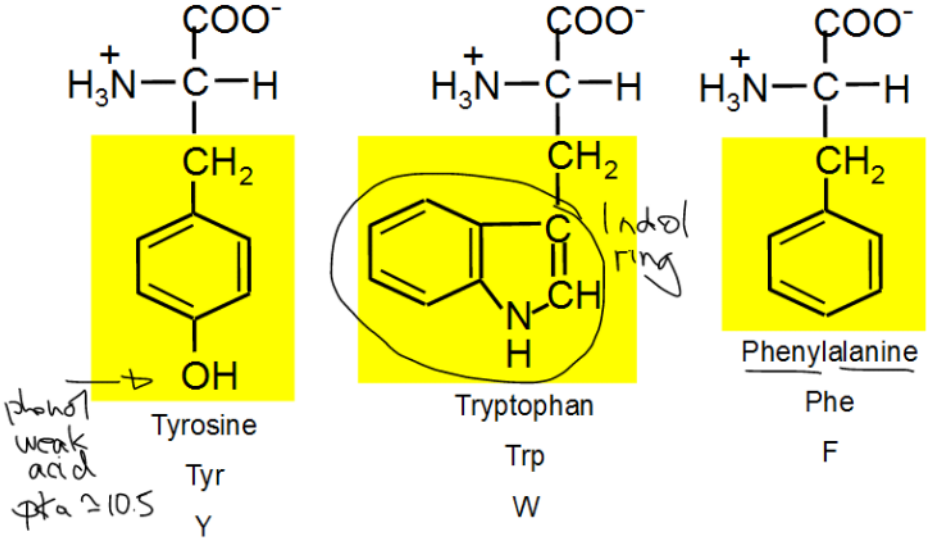


    - Proline contains cyclic secondary amine



    - Central carbon on the backbone is the alpha carbon, beta carbon is the first carbon on the sidechain
    - Isoleucine is a constitutional isomer of Leucine - beta carbon is also a stereogenic centre (2 STEREOGENIC CENTRES)
        - Beta-carbon is always in D conformation (L conformation is Aloleucine)

Aromatic nonpolar amino acids: Tyr, Trp and Phe

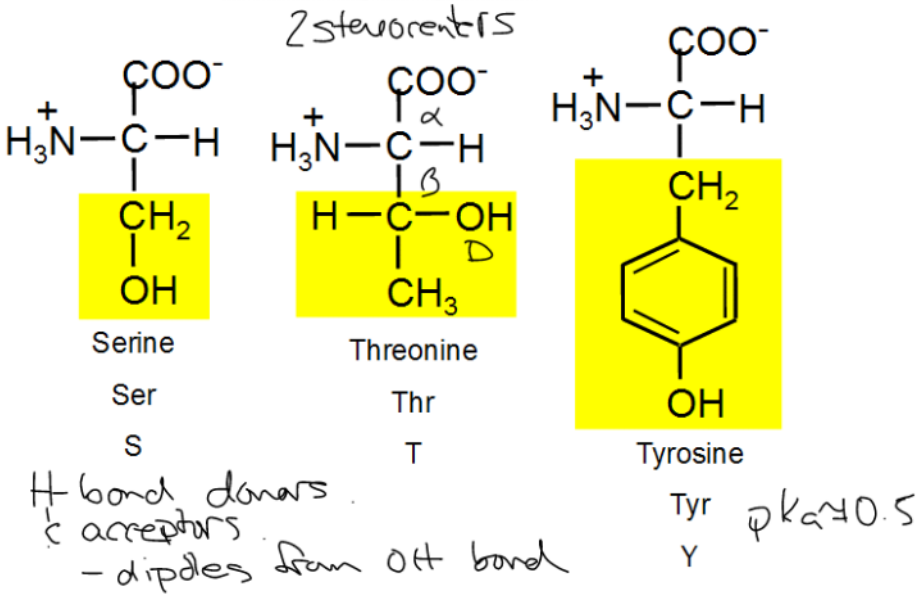


- Tyrosine
  - Phenol are weak acids - Tyrosine pKa approx. 10.5
  - Tyrosine can be deprotonated - below pH 9.5 it is fully protonated
- Tryptophan
  - Indol ring - 6 membred ring fused to a five membered ring contain N and the whole system is aromatic - non polar due to aromaticity
  - Least used amino acid - usually has a specific function
- Phenylalanine
  - Benzene attached to an alanine

Non polar Amino acids

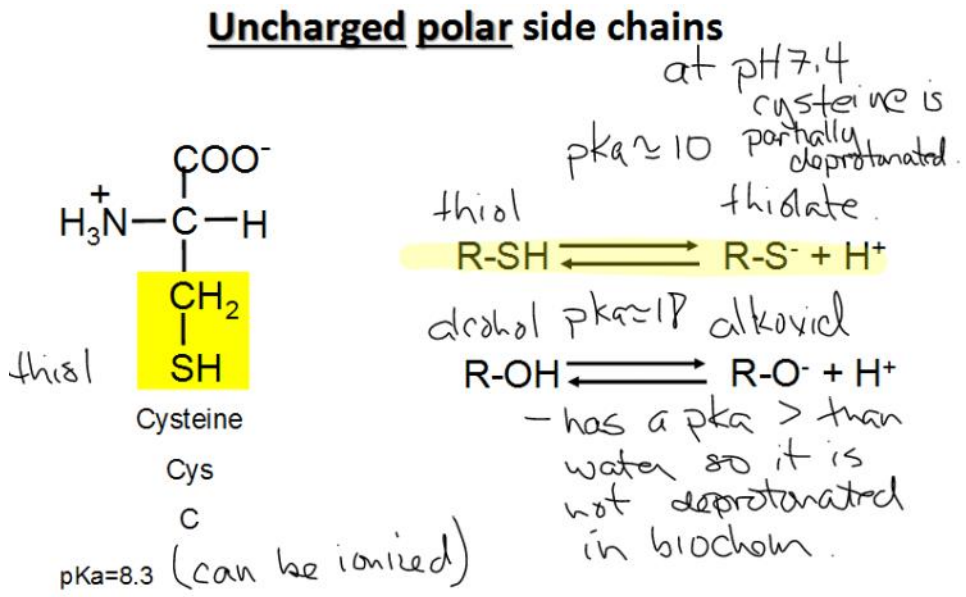
- All amino acids have alkyl R groups (Ala, Val, Leu, Ile)
- Tyr can behave polar or charged (depending on pH)
- Generally pretty hydrophobic - exception Pro, Gly, Tyr and Trp

Uncharged polar side chains



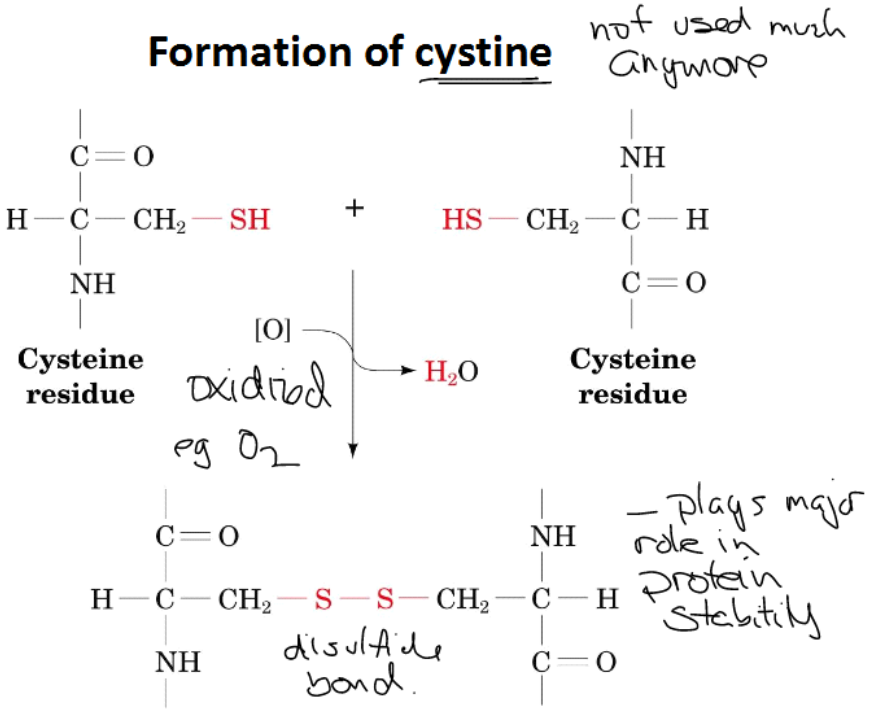
- Typically contain Oxygen - high EN - any protons attached the bond will be highly polarized

- All contain H-bond donor and H-bond acceptor (2 lone pairs on side chain O)
- Threonine - beta branched - containing a stereogenic centre (2 STEREOGENIC CENTRES) - must be D conformation on the beta carbon



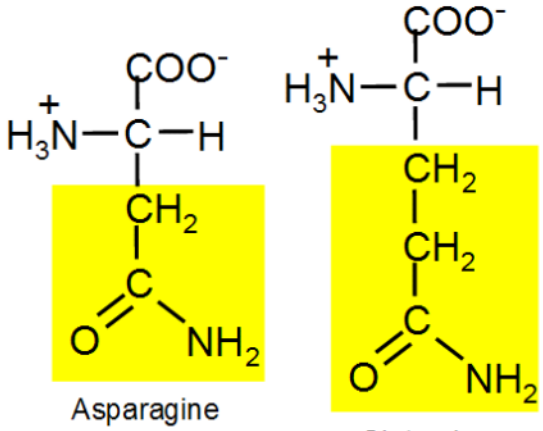
- Thiol (SH) works as a nucleophile, Acetyl COA, Cysteine are very important

- Thiol (SH) works as a nucleophile, Acetyl COA, Cysteine are very important
- Can be deprotonate - thiol pKa =8.3
- Thiol deprotonates to a thiolate, Alcohols deprotonate to an alkoxide
- Alcohol pKa = 18
- Sulfur can stabilize negative charge, interact with solvent better so it is much easier to take protons off
- Serine and Threonine pka is above 16 so it does not deprotonate in biochemistry
- Thiols are readily deprotonated - at pH 7.4 -partial deprotonation (about 10%)



- Thiols are very senstive to oxidation - O2 in the air can react O2 into water to make a disulfide bond
- Product called cystine or disulfide of cysteine
- Cross-link in proteins - inportant 3D shape and function of many proteins

Uncharged polar side chains



Asparagine

Asn

N

Glutamine

Gln

Q

After by 1 CH_2

amide side chains

H bond donors & acceptors

- Won't get deprotonated - due to amide resonance
- H - bond donors and acceptor

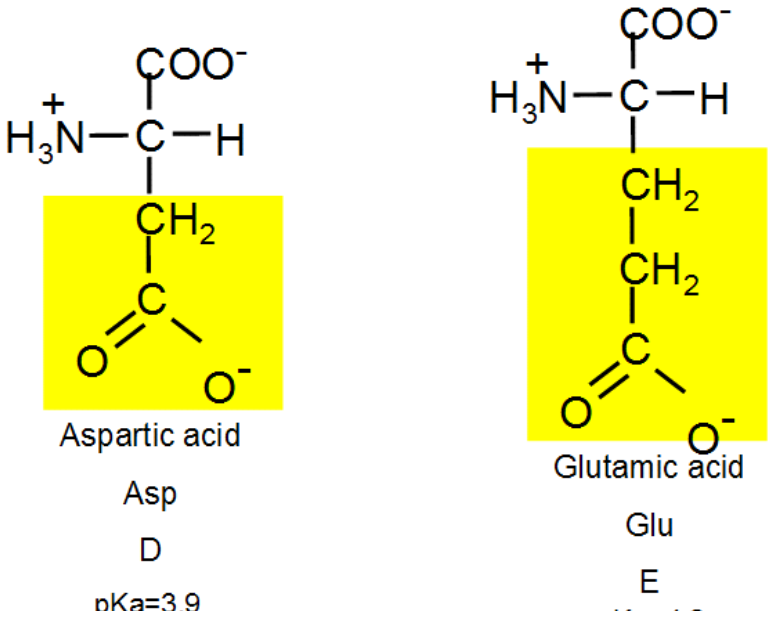
Uncharge polar side chains

- Can form hydrogen bonds with water - polarized bonds so can intereact wtith dipole of water
- Alcohols in R-groups (Ser, Thy, Tyr)
- Tyr contains phenol that can be deprotonated
- Amide groups - interact readly with water
- Usually more soluble, except Tyr
- Thiol group: Cys
  - Very good nucleophile

19/01/2017 8:20 PM

- Read chapter 4 for amino acids
- Chapter 3 for nucleic acids

Charged polar (acidic) side chains



Aspartic acid

Asp

D

pKa=3.9

Glutamic acid

Glu

E

D

pKa=3.9

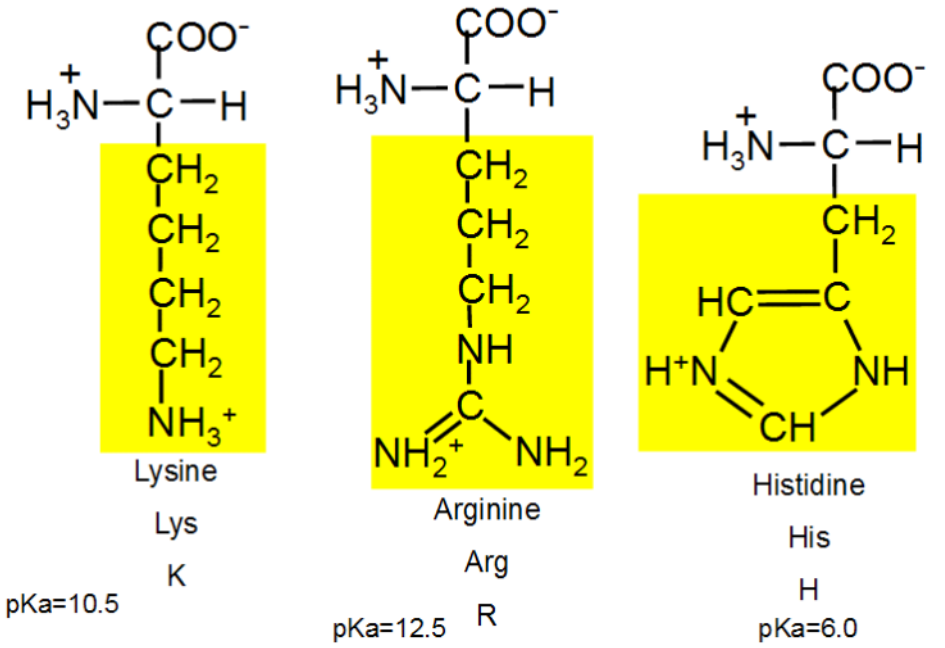
Glu

E

pKa=4.2

- Carboxylic acid sidechains - differ by CH2 group
- Just report both as pKa = 4
- Charged amino acid side chains have a charge
- Net negative charge at pH 7 - deprotonated
  - Carboxyl group - alpha pKa = 2.2 (deprotonated)
  - Side chain pka = 4.0 (deprotonated)
  - Alpha amino group = 9.4 (protonated)
  - Play a role in metal binding sites
  - Moderately good nucleophiles and electrostatic interactions (salt bridges)

Charged polar (basic) side chains

- 

- 1
- Lysine
  - Protonated at pH = 7.4
  - Neutral form is an amine and protonated - called ammonium, called the epsilon amino group - the carbon attached to the ammonium
- Arginine
  - Guanidinium in protonated form, guanidine if deprotonate (neutral)
- Histidine
  - Aromatic amino acid
  - Imidazol - lone pair from the nitrogen creates aromaticity
  - Pronated called Histidinium and the imidazol called imidazolium
  - Frequently protonated or deprotonated due to being close to physiological pH

Basic Amino Acids

- Net positive charge at pH 7 - Lys and Arg fully protonated at pH 7
  - Participate in electrostatic interactions -salt bridges (mainly Lys, Arg)
- His has a side chain of pKa 6.0 so 10% pronated at pH 7
- His can act as both proton donor and base
- His contain proteins are biological buffers - major role in catalysis

Properties and Conventions Associated with the Standard Amino Acids

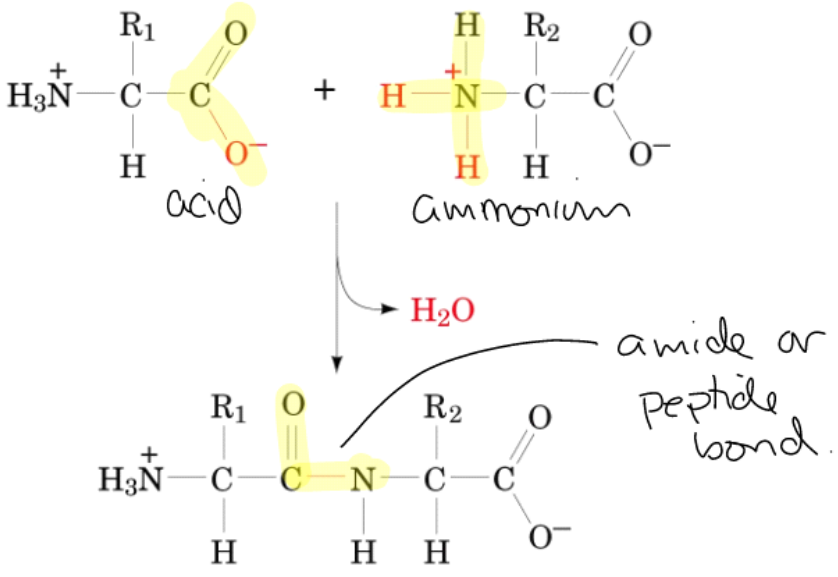
Amino acid	Abbreviated names		M_r	pK_a values: pK_1 (—COOH)	pK_2 ($—NH_3^+$)	pK_R (R group)	pI	Hydropathy index[a]	Occurrence in proteins (%)[b]
Nonpolar, aliphatic R groups									
Glycine	Gly	G	75	2.34	9.60		5.97	−0.4	7.2
Alanine	Ala	A	89	2.34	9.69		6.01	1.8	7.8
Valine	Val	V	117	2.32	9.62		5.97	4.2	6.6
Leucine	Leu	L	131	2.36	9.60		5.98	3.8	9.1
Isoleucine	Ile	I	131	2.36	9.68		6.02	4.5	5.3
Methionine	Met	M	149	2.28	9.21		5.74	1.9	2.3
Aromatic R groups									
Phenylalanine	Phe	F	165	1.83	9.13		5.48	2.8	3.9
Tyrosine	Tyr	Y	181	2.20	9.11	10.07	5.66	−1.3	3.2
Tryptophan	Trp	W	204	2.38	9.39		5.89	−0.9	1.4
Polar, uncharged R groups									
Serine	Ser	S	105	2.21	9.15		5.68	−0.8	6.8
Proline	Pro	P	115	1.99	10.96		6.48	1.6	5.2
Threonine	Thr	T	119	2.11	9.62		5.87	−0.7	5.9
Cysteine	Cys	C	121	1.96	10.28	8.18	5.07	2.5	1.9
Asparagine	Asn	N	132	2.02	8.80		5.41	−3.5	4.3
Glutamine	Gln	Q	146	2.17	9.13		5.65	−3.5	4.2
Positively charged R groups									
Lysine	Lys	K	146	2.18	8.95	10.53	9.74	−3.9	5.9
Histidine	His	H	155	1.82	9.17	6.00	7.59	−3.2	2.3
Arginine	Arg	R	174	2.17	9.04	12.48	10.76	−4.5	5.1
Negatively charged R groups									
Aspartate	Asp	D	133	1.88	9.60	3.65	2.77	−3.5	5.3
Glutamate	Glu	E	147	2.19	9.67	4.25	3.22	−3.5	6.3

- Alpha carboxyl group - pKa 2.2
- Alpha amino group - pKa 9.4
- 7 amino acids with relevant pKa (plus or minus 0.5)
    - Tyr - 10, Cys - 8, Lys - 10.5, His - 6, Arg - 12.5, Asp - 4, Glu - 4
- Know the structure, include stereochemistry, be familiar with their name, 3 letter code, 1 letter code, pKa of all ionizable groups and if they are charged or not charged given conditions

Amino acids can join via peptide bonds

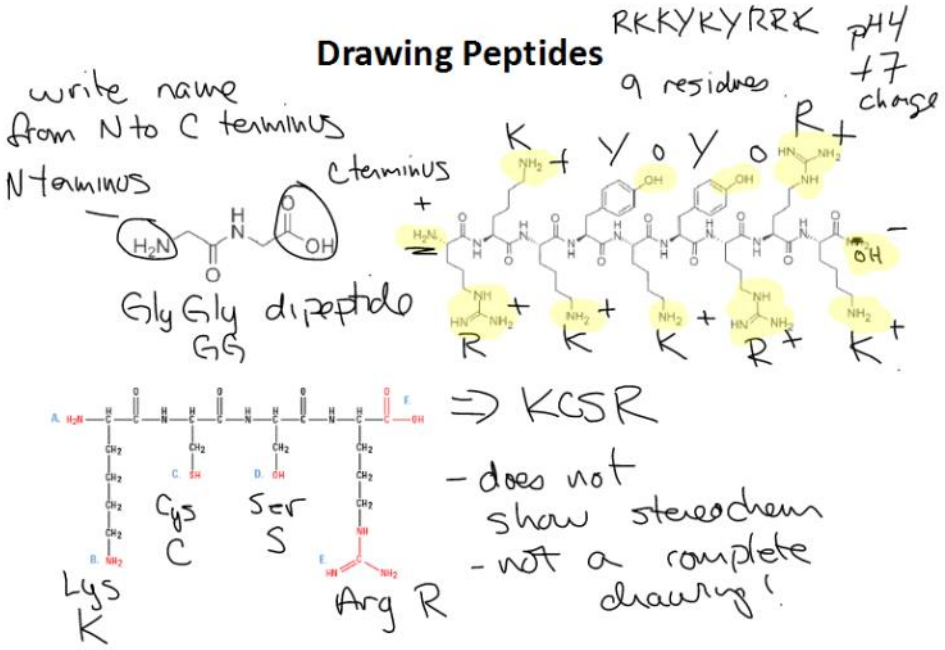
- The amino and carboxyl group allow amino acids to polymerize
    - Reat haid to teail fashisn
        - Elimiation of water molecule
        - Formation of a colavent amide bonde called a peptide bond
    - Peptides
        - 2 amino acid rsidues - dipeptide
        - A few aa peptide - oligopeptide
        - Many - polypeptide

Condensation of two amino acids to form peptide bond



- Form an amide/peptide bond - ammonium and carboxylate cannot react in this form,

they must be activated in a ribisome, electrophile activated



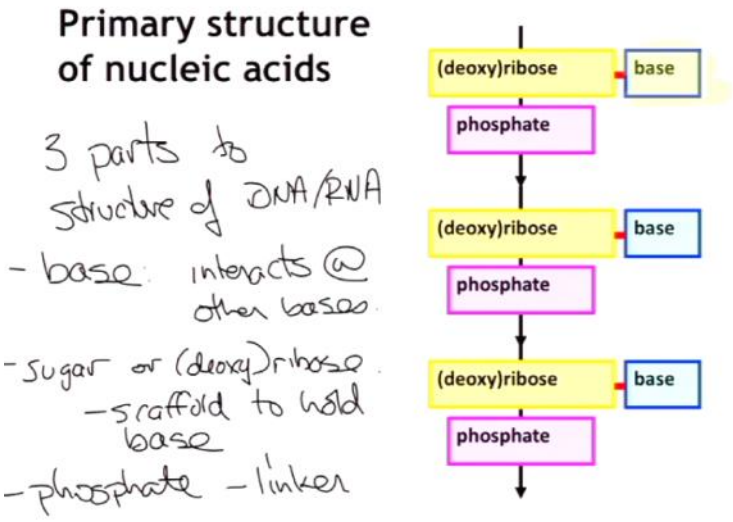
- Directionality - convention for writing peptide sequence
  - Start from the N terminus from the left to the right to the C terminus
  - If pH is stated - the charge should be shown, without it, show the neutral
  - form

Nucleic Acids - Lecture 4 and Lecture 5

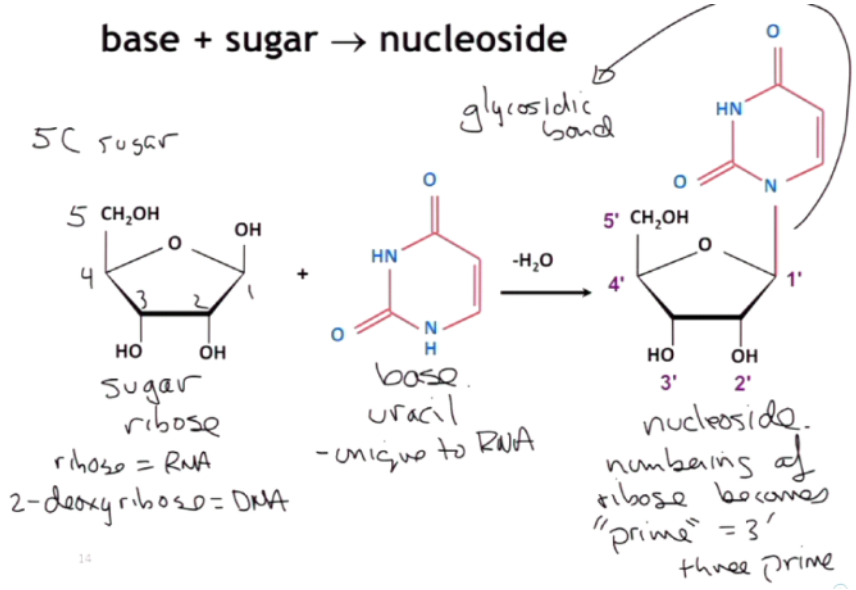
January-19-17 8:20 PM

Nucleic Acids

- Not a single building block - composed of 3 parts
  - DNA and RNA
    - Base comoponent - interacts with other bases - information storage and convey info
    - Sugar or (deoxy)ribose - scaffold to hold base, allow the different monomers to be linked together
    - Phosphate - linker - links ribose together
    -



- Base + sugar is a nucleoside
  - Held together by a glycosylic bond (bond unqiue to sugar)
  - Sugar is always 5 membered
    - In Ribose - Alcohol on 2C and Alcohol on 3C are always down
    - In DNA - 2- deoxyribose and in RNA - ribose
  - When the ribose becomes bonded, numbering changes from 1 to 5 to 1' to 5'



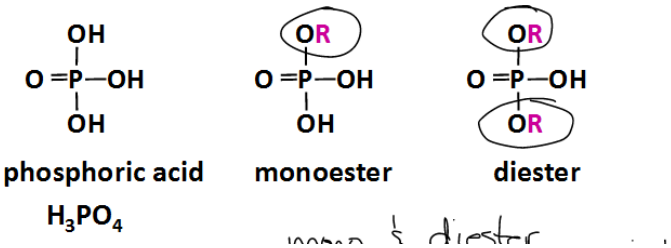
- Glycosidic bond - links the base to the ribose
  - Called this because it occurs as the C of the sugar that has 2 heteroatoms bonded to it (O-C-N)
  - Up or down - up is Beta in nucleosides, always the Beta conformation

Phosphate Esters

- Connect Nucleosides together
- Phosphoric acid has 3 protons - take one H away and react with an alcohol we get a phosphate ester (Monoester), repeat the reaction to get a diester
- Monoester and diester is common and triester do not occur in biochemistry - stable bonds
- The fact we can forma diester- we can link two sugars together, reacting at -OH sites

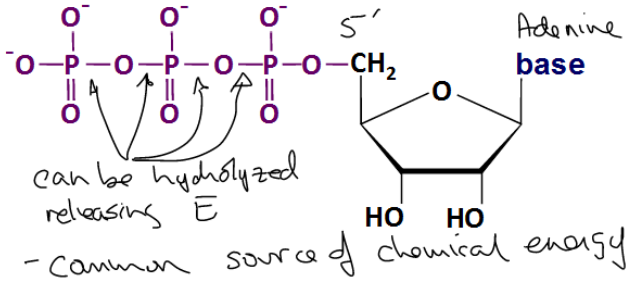
phosphate esters

linker - connects nucleosides together



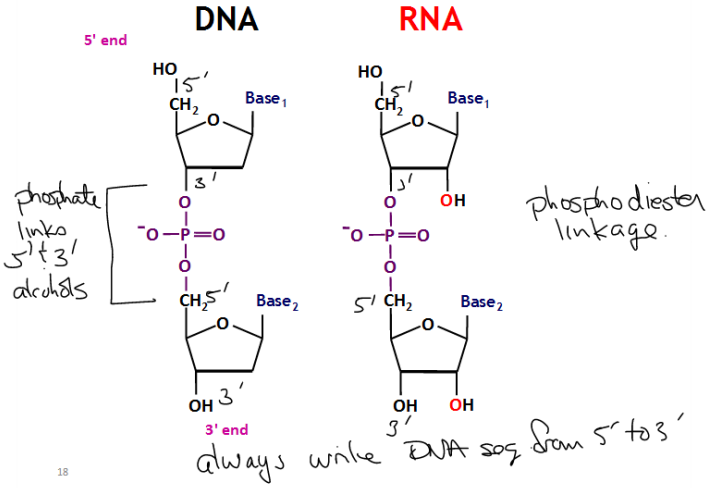
Nucleotide - base + sugar + phosphate
- Link shown at the 5'

Nucleotides
- Used to store energy as ATP, high-energy nucleotides
  - ATP adenosine 5'-triphosphate



- At the end - monoester, -30 kj/mol, high energy bond

DNA and RNA strands

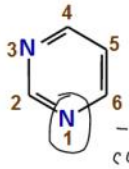


- Phosphate diester links the 5' and 3' alcohols - bond known as a phosphodiester linkage
- Write down DNA sequence from the 5' sequence to the 3' sequence

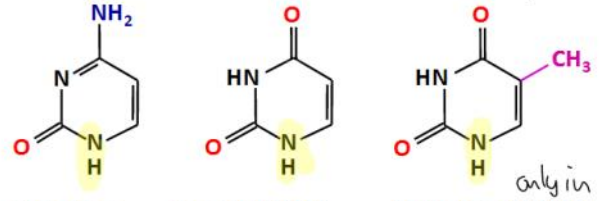
The Nucleic acid bases

Pyrimidines

3 pyrimidines in DNA/RNA



- attaches to 1' carbon of sugar



cytosine — C — DNA/RNA

uracil (RNA) — U only

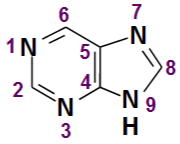
thymine (DNA) — only in — *(5-methyluracil)* — T

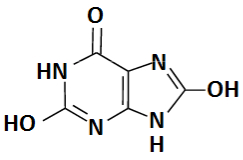
- All nuclear bases are aromatic - double bonds or a lone pair of electrons on a Nitrogen
- NH - H bond donor, O - H bond acceptor

2nd type of base

- 2 aromatic rings

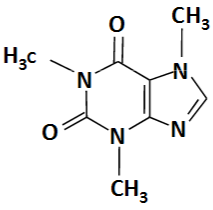
Purines





uric acid

- birds & reptiles excrete nitrogen via uric acid

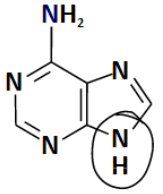


caffeine

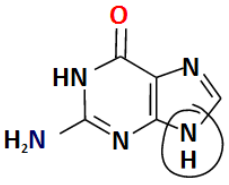
= can play roles in biology apart from DNA/RNA

- No N-N bonds, all aromatic rings, birds and reptiles excrete nitrogen via uric acid
- Caffeine -

2 purines in DNA/RNA



adenine — A

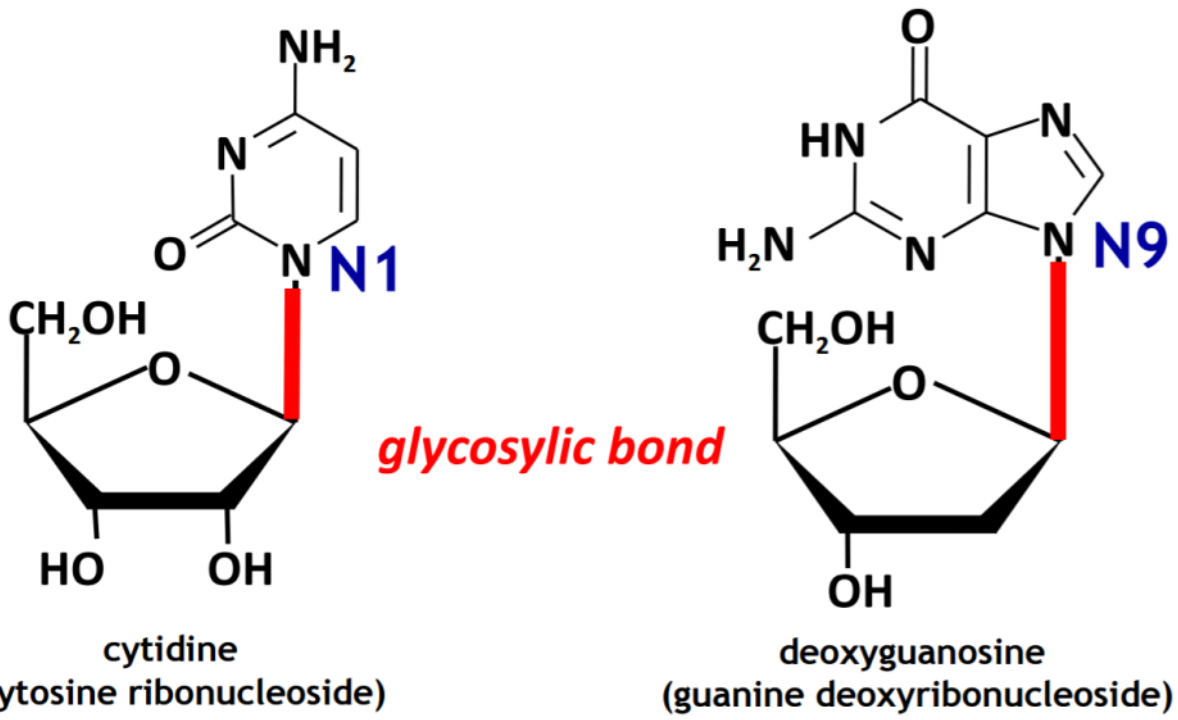


guanine — G

— linked to sugar C1' by N9

- note no N-N bonds!

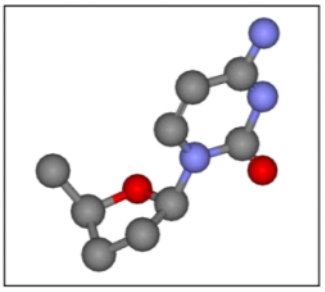
- Read into Chapter 3 (Part 1 and 2)
- Read Chapter 8 for Carbs / lipids
- All prymidines are attached to the sugar via NI
- Both purines attached via N9 sugar
- 2' Oxy - Ribose (RNA) 2' deoxy -DNA



cytidine
(cytosine ribonucleoside)

deoxyguanosine
(guanine deoxyribonucleoside)

- Single bonded (glycosylic bond) - could have many configurations but that is not the case due to steric hindrance which limits roatation about he gloycolsylic bond
- This allows H bond acceptors and donors to be pushed away from the sugar allowing for it to interact with other bases
  - WITHOUT steric hindrance there would be no base pairing



syn *anti*

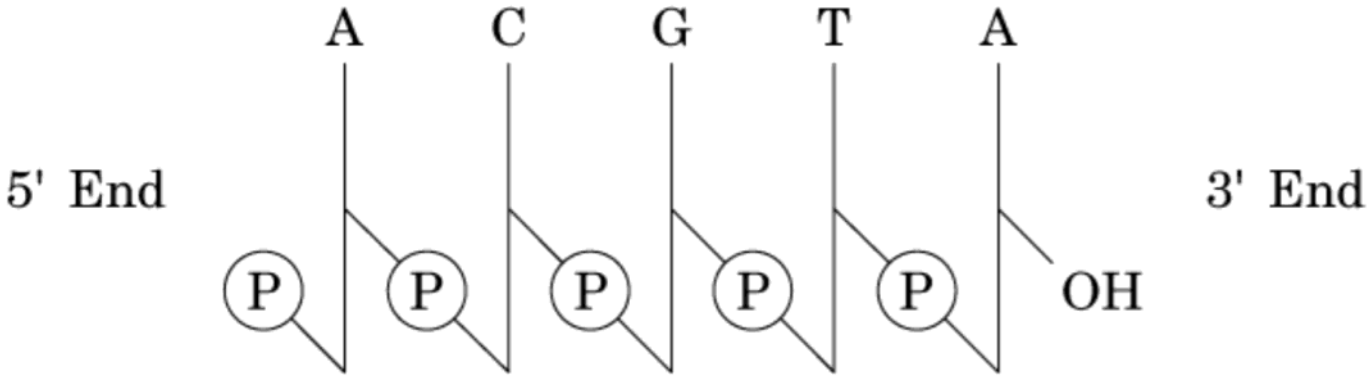
Base + sugar → nucleoside

.

Base	Nucleoside
Adenine	Adenosine
Guanine	Guanosine
Cytosine	Cytidine
Thymine	Thymidine
Uracil	Uridine

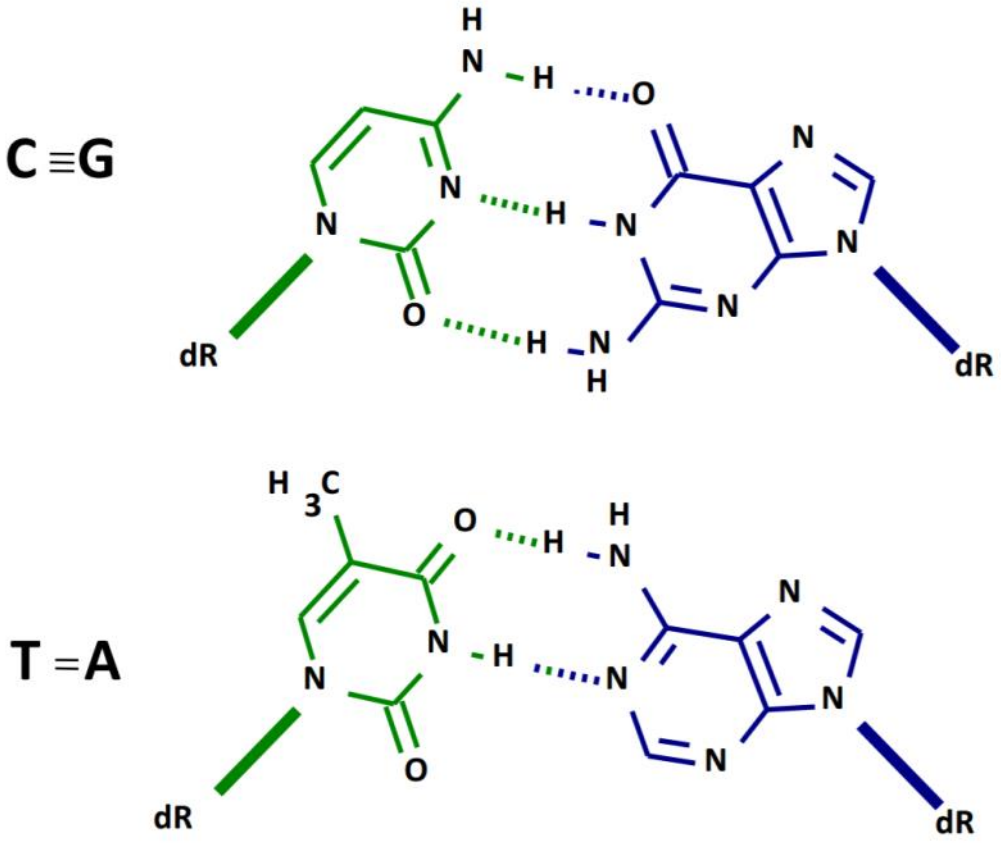
- All of this can be DNA or RNA EXCEPT Uracil which is only RNA and Thymine which is only DNA
- When A is attached to the ribose, it is said as A. If attached to deoxyribose, it is said as dA. ATP is on ribose.

Nucleotide sequence written from 5' to 3'



5' ACGTA 3'

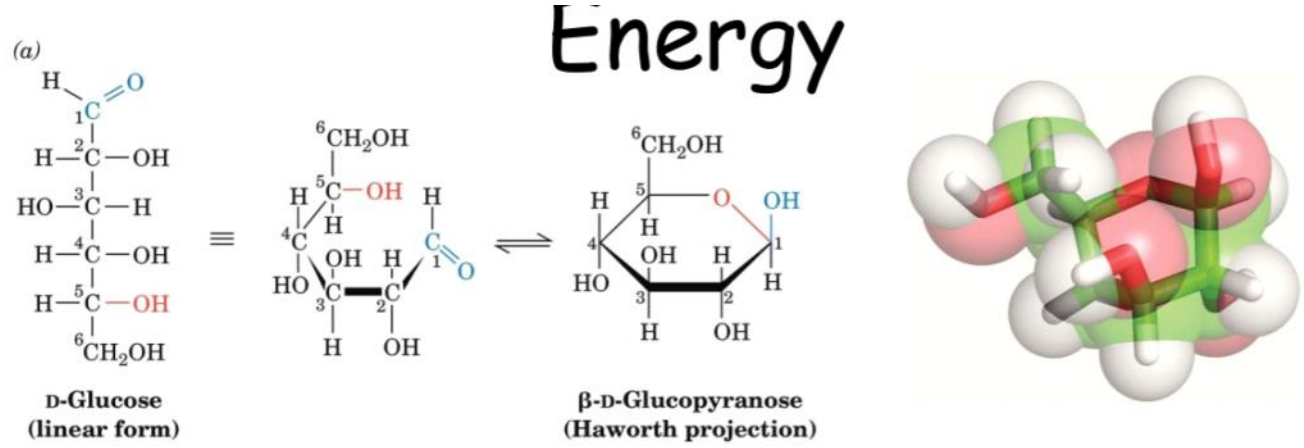
- Watson crick base pairs for DNA - allows for them to be complementary to template their replication
- All driven through hydrogen bonding. Dictates base pairing, forms 3H bonds - all of these bonds are linear
  - CAN BE EXPLAINED BY ORBTIAL APPROACH - very strong H-bonds]
- Strain free, highly stable structure
- GC has a stronger interaction (3H bonds versus 2H bonds) - PCR amplification can be difficult if GC rich
- C can only H bond with G

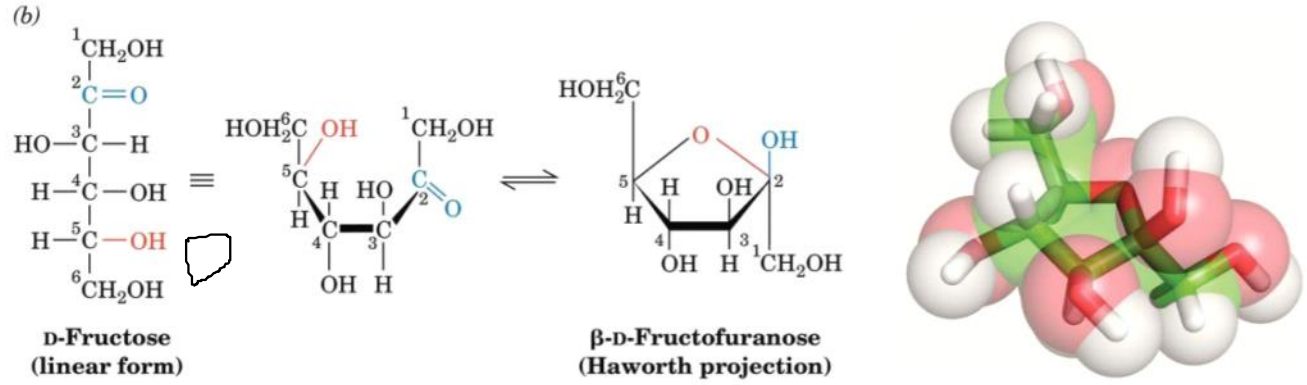

C ≡ G
H
N
H
O
N
N
H
N
N
N
O
H
N
H
dR
dR
T = A
H 3 C
O
H
H
N
N
N
H
N
N
O
N
dR
dR

Carbohydrates - Lecture 5

January 24, 2017 1:01 PM

- Can also be found covalently bonded to proteins through serine, theornine or asaparagine - glycoproteins
- Glycobiology - study of carbohydrates in biology
- Carbohydrates have a genral formula of $C_xH_{2x}O_x$





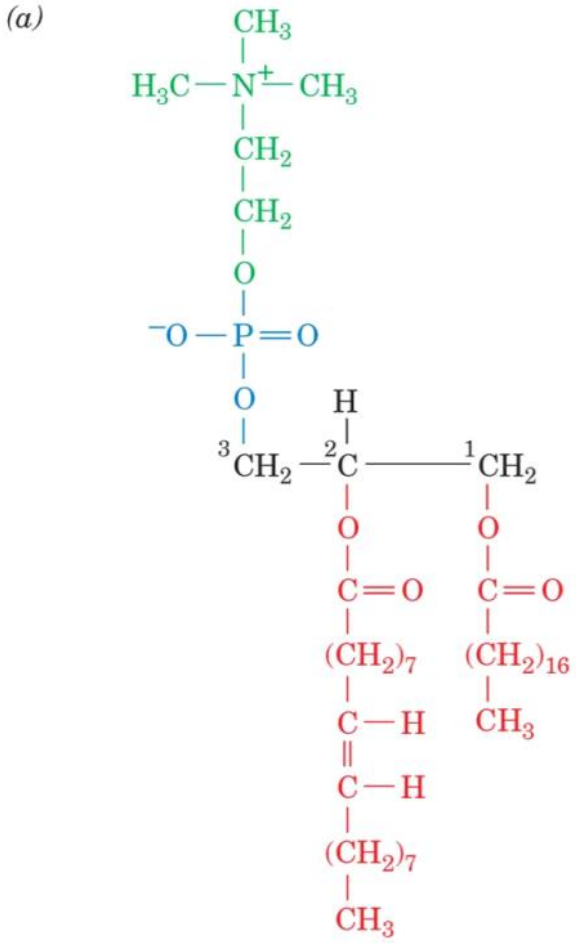
- Can form a hemi-acetal to react the aldehyde and alcohol, called an aldose (glucose), fructose is called a ketose
- Sugars are usually drawn in the Haworth projection (cyclic), fisher projection is the linear form
- Both can be used for energy

- Polysaccharides
  - Dissacharide - 2 sugars
  - Glycosylic bond at anomeric carbon, can be a source of energy
  - Lactose is usually made
  - Glycogen, sucrose are also sources of energy
  - When polysaccharides get really big, their properties change, over 1000 sugar resides - very large polymers
    - Can self assemble, H bond the chains together - powerful forces
    - Fibrous, rigid and isoluble - ex. Cellulose - structural

Lipids - Chapter 5

January 24, 2017 1:40 PM

- 4th class of biopolymers
  - Less complex than other polymers
  - Formation of membranes
- Fatty acids are typically between 12 and 22 in length
- A long carbon chain (unsaturaed or saturated) bond to a carbonxylic group is a fatty acid
  - Saturated - no double bonds
  - Unsaturated - 1 -6 double bonds
- Unsaturation
  - Introduces a double bond - flat and planar, putting a big bend in the chain
    - Impact in the fatty acids - hard to pack them together
    - Lose a lot of London forces as things are further apart
      - ☐ Can be seen by melting point saturated versus unsaturated fatty acid
      - ☐ Mp. Of Stearic acid - 69.7°C, oleic is 12°C
    - Use a mixture of fatty acids to control the fluidity of supramolecular lipid structure
- Enrgy storage
  - Triacylglyerides
    - 3 fatty acids on a glycerol backbone
    - Broken down to Acetyl-Coa and make ATP - stored energy
  - Membranes
    - Similar to triacylglycerides - one fatty acid changed by a phosphate
    - AMPHIPHILIC - charged polar head group and non-polar tail
  - Phosphate group is modified by a head group (Ethanolamine, Choline, Serine)



1-Stearoyl-2-oleoyl-3-phosphatidylcholine

- Depending on the temperature and conditions, it can control the fluidity using saturated fatty acids

MIDTERM 1 REVIEW

January 24, 2017 1:57 PM

- 22 questions
  - Some - multiple choice - might be 2 or 3 correct answers
  - True/False
  - Short answer
    - Peptide sequence - charge at the pH given
  - Long answer
    - More space to write an answer in
  - Structure of the 2 carbohydrates, ribose
  - pH, pKA, buffer composition, intramolecular forces

3D Structure of Proteins - Lecture 6

January 31, 2017 1:03 PM

- Primary structure - order of the amino acids making up the polypeptide
  - Typically 300 amino acids
- Secondary structure - local structures formed based on the sequence
  - Only subsections of the primary structure
  - 6-24 residues
- Tertiary structure - everything from the primary structure forms the tertiary
  - Folded protein domains are made
- Quartenary structure - several tertiary structure interacting together
  - Multiple folded domains assemble to make complexes

- Structure of Proteins
  - Will adopt a specific 3D structure (3D conformation)
    - Unusual - most organic polymers do not have the same 3D structure - polystyrene (totally random)
  - The 3D structure is called the native fold
  - Structue is crucial to the FUNCTION of the protein
  - The Native old has a large number of favourable interaction within the protein
    - Protein folding is costly - the primary structure has a chaotic number of conformations
    - By creating a 3D structure, we reduce possible states to a single one - entropy is significantly reduced
    - Entropy cost is reduced by enthalpy - the polypeptide is stabilized in the process of protein folding (favourable interactions)

- Weak non-covalent interactions
  - Hydrophobic effects
    - Decreases solvation shell around the hydrophobic groups - pushes hydrophobic groups together
    - Val, Ile, Leu, Met, Phe - side chains on interior of protein
  - Hydrogen bondings
    - Backbone amide and carbonyl groups (Proline does not have this)
    - Between side chains - interactions Ser, Thr, Asn, Gln, Trp, Tyr
  - London dispersion
    - Fairly weak - but several group interacting it can be stronger
    - Groups pack closely on interior of proteins (packing density)
      - ☐ Packing density - oil drop 0.6, crystal - 0.7, proteins - 0.75
      - ☐ The high packing density is due othe side groups in the interior being very close together, strengthens london dispersion forces
  - Electrostatic interactions
    - Charged residues can interact if close together - salt -bridge
    - Involved in the inside of proteins - not weakened by solvation shells
    - On the outside, they work to attract ions and charges towards each other
- The Peptide bond
  - Plays a role in the way residues interact
  - Amide linkage
    - Partial double bond character
    - Shorter that a regular C-N single bond (approx. 40% double bond character) - 0.13 A shorter
    - Longer than a regular C-O double bond - approx. 0.02 A longer
    - Results in restricted rotation - the amide is always planar
      - ☐ Therefore there is a few number of rotatable bonds
- Two potential stereochemical nature
  - Cis and Trans - IN BIOCHEMISTRY
    - IN BIOCHEMISTRY - we are interesting in the location of the CARBONS

            - ☐ Carbon chain is TRANS across the amide bond
            - ☐ Carbon chain is CIS across the amide bond
    - Cis configuration has destabilzing steric interactions - unfavourable (only 1% of amide bonds in this configuration)
        - TRANS is most common
    - Proline is unique - can adopt both
        - Both trans and cis have steric interactions - so both are equally likely to occur
        - Can much more easily adopt trans and cis bonds
            - ☐ Potent in forming turns in a primary sequence (beta-turns)
            - ☐ 10% in the cis form
- Peptide bonds are planar
    - Dihedral angle - angle between 4 points
        - 4 atoms and 3 bonds
        - Rotation state of the bond between pair of atoms
        - Look at It like a newman projection
            - ☐ The angle between the two pairs is the dihedral angle
    - The dihedral angle is called the omega angle
        - Almost always 180^{o} for almost all amides
        - In Cis, omega angle 0^{o}
- Peptide plates
    - Rigid, flat amide bonds
    - The flat amide group is connected to two rotatable carbons
    - Amino group to the alpha carbon and the alpha carbon to the carbonyl group - both from dihedral angles
    - 2 Dihedral angles for EACH alpha carbon
        - Phi - alpha carbon - amino bond
        - Psi - alpha carbon - carbonyl bond
    - Each Phi and Psi angles defined the structure of the protein as they describe how peptide plates and side groups interact
    - Most fully extended bond - Phi and Psi equal to 180^{o}
- Many angles are prohibitied due to steric overlap
    - Major steric clash if 2 cis phi and psi angles adjacent together
    - Forbidden angles
        - Can be calculated
        - Figure to be that most angles cannot be adopted
    - Polyalanine was used - loses symmetry due to a stereocentre
        - True for all non-beta-branched amino acids)
        - Beta-branched aa (Val, Thr, Ile, Leu) have smaller ranges of allowable angles
    - Glycine has very little steric hindrance
        - Ton of available Phi - Psi angle space to adopt
        - No side chain - removes side chain steric interactions
        - Important for making sharp turns
    - Proline - most rigid aa due to the ring
        - Phi angle is severely hindered
        - Phi angle stuck at -60^{o} in a range of 25^{o}
- In real proteins - some Phi Psi angles are in forbidden space
    - Omega angle can sometimes undergo a twist (between carbonyl and nitrogen - amide bond)
        - This enables a forbidden angle to be populated
- Ramachandran plots define the secondary structures that form
    - 3 areas of allowable angles
- Secondary structures
    - All structures are different amounts of
        - Alpha helix
        - Beta sheet
        - Turns or bends
            - ☐ Globular
                - ◆ 30% alpha helix

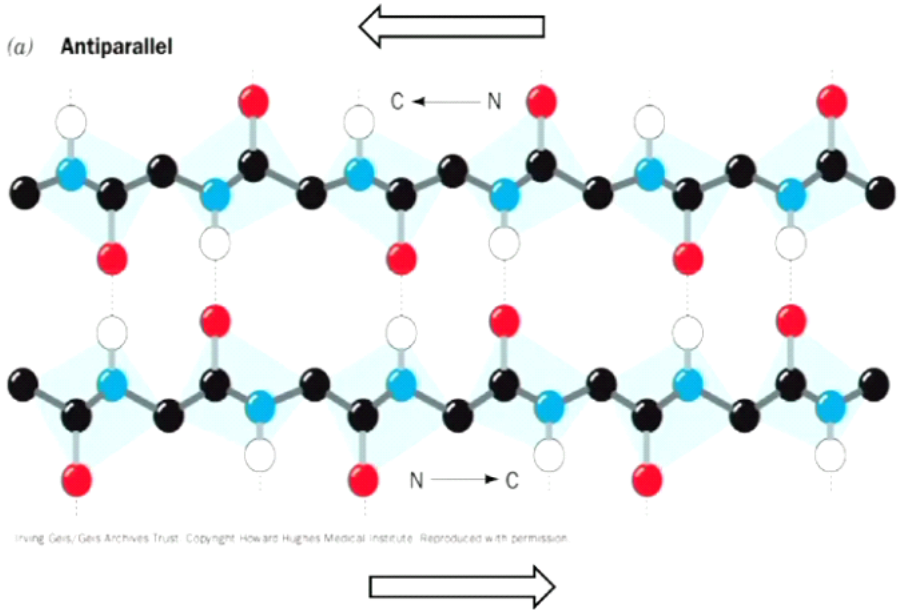
        - 28% beta helix
    - Myglobin
        - ALL alpha helices
    - Immunoglobulin
        - All Beta-sheet
- Alpha helix (helical structure)
    - Pitch
        - How much it goes up or down on each turn
        - Number of residues it takes to go around the helix
        - Handedness - upwards right handed or left handed position
        - N = residues/turn
        - P = rise of the helix/turn
        - Residue = amino acid/peptide subunit
- Beta helix

Lecture 7 -

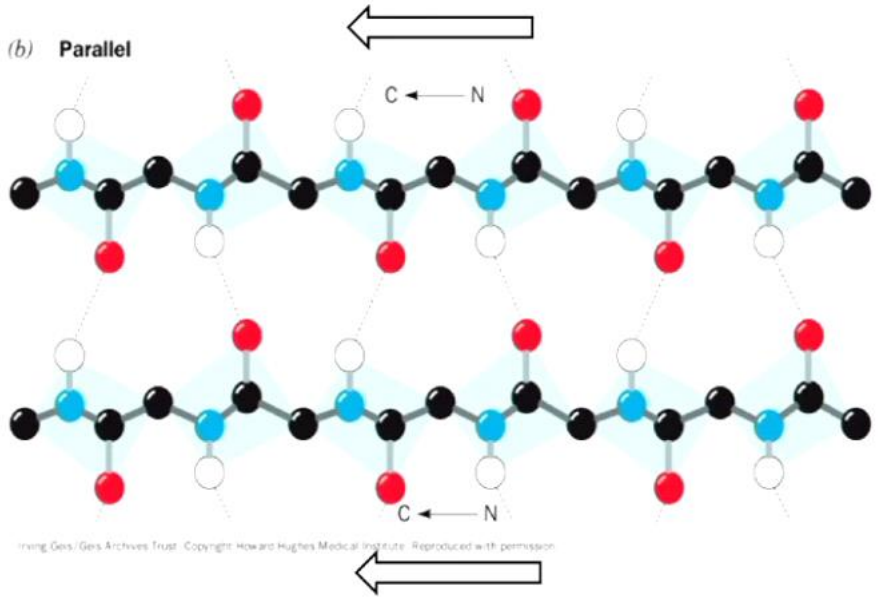
February 7, 2017 2:54 PM

- Right handedness or left handedness
- n above 0 - right handed
- n below 0 - left handed
- Archetypical alpha-helix
  - *n* = 3.6 residues/turn
  - p = 5.4 A
  - Phi = -57º
  - Psy= -47º
- Range from 12 to 140 amino acids
  - Typically 20 residues
- Very common structures in proteins
- Backbone H bonds
  - NH of both reside to C=O of (n-4)th residue
  - Linear H bonds - very strong H bonds
  - The side chain is pointing away from the core of the helix
  - Tight packing of inner core
    - Maximize London forces
- Helical structures
  - Nm
  - n = number of residue per helical turn
  - m = number of atoms, including H, in the ring that is closed by the hydrogen bond
    - H-bonds stabilize helix, important in describing the structure
    - Hydrogen bonding pattern of several polypeptide helices
      - ☐ 3_{10} tighter and thinner - common at the end of a alpha helix
      - ☐ The most stable is an alpha helix
      - ☐ Pi helix - wide, flat helices - m = 16
      - ☐ 2.2_7 has not yet been observed
  - 3_{10} - n = 3 p=6A
  - Alpha n=3.6 p=5.A
  - Pi n=4.4 p=5.2A
    - Greater interior space - destabilizes it
- Polyproline II helix
  - Missing NH - no h bonding
  - Locked Phi angle difficult to fit into alpha helix
  - n=3 p=9.4A
    - Very extended
  - Glycine can also adopt poly proline III helix
  - Pro -Gly - extended helices found in collagen
    - Extremely proline glycine rich protein
    - Essential part of its structure
  - Conformational difference - change by rotating bonds
  - Kept in a helix because it is the most stable conformation to prevent steric interaction
- Beta structures
  - Beta pleated sheets H bonding occurs between neighboring strands
  - 2 types
    - Antiparallel B pleated sheet
    - Parallel B pleated sheet
      - ☐ 2-22 strands average 6 strands each can have up to 15 residues
      - ☐ Typically approx. 6 residues

- The antiparallel β pleated sheets.

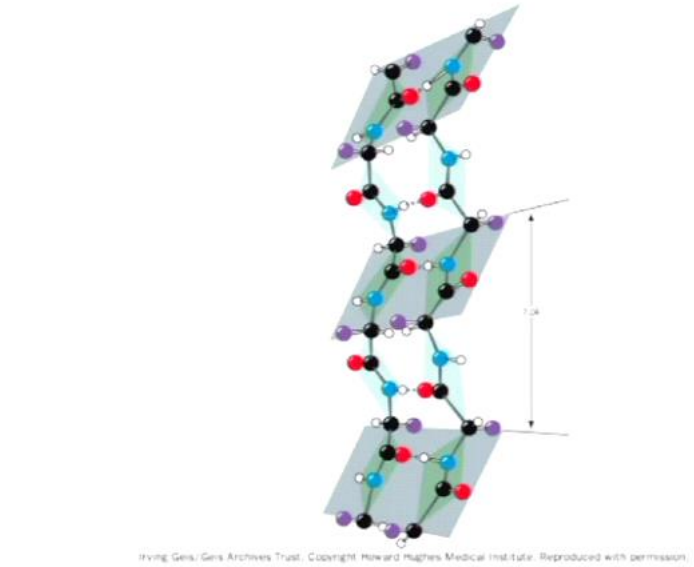


  - Sites of H bonding
    - Linear - strong
- The parallel β pleated sheets.



  - Not parallel
    - Decreased strength in each bond
    - Less stable than antiparallel
  - What does pleated mean?

“pleated” two-stranded antiparallel β sheet



Irving Geis/ Geis Archives Trust. Copyright Howard Hughes Medical Institute. Reproduced with permission.

- Bent at the alpha carbons - all bend points are alpha carbons
- N to n+2 are point in same direction, n to N +1 - opposite directions
- Side chains point away from sheets

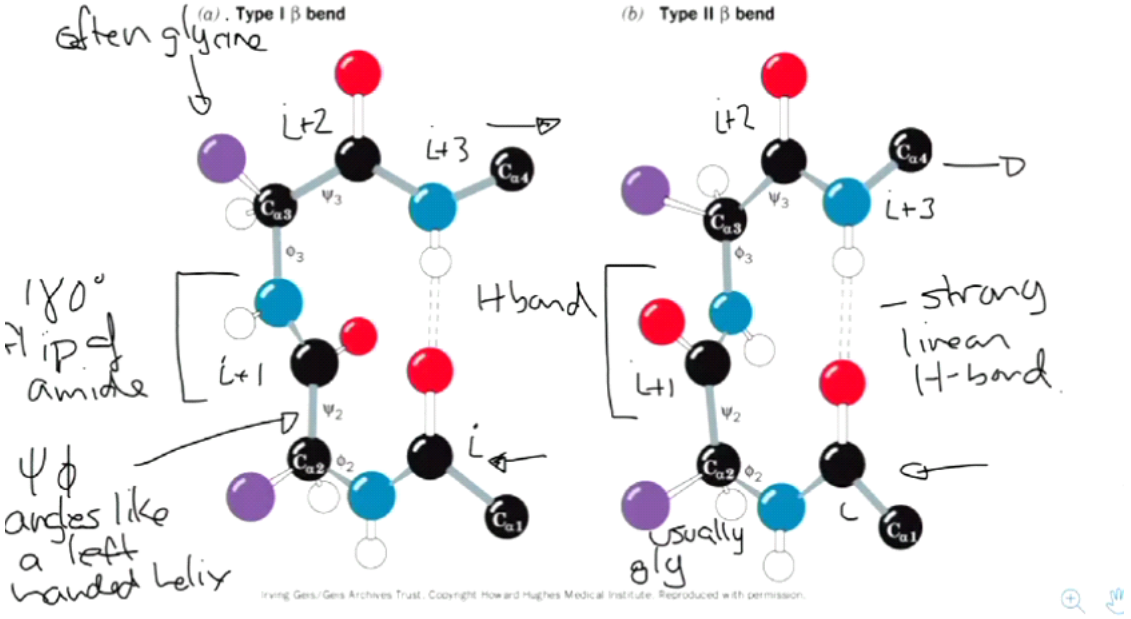
- Coil and Loop conformations
  - Random coil - non structured - monomers are orientated randomly - no specific shape
  - A turn
  - 40% of protein sequence Is coil and loop conformations
  - Globular protein are largely 2 structures joined by abrupt changes
  - Connect strand of antiparallel Beta sheet - TIGHT TURN
  - Occur on surface of protein
    - Alpha and beta structure in the interior
  - 4 successive residues arranged in one of two ways
- Beta Hairpins
  - Type I bond
    - I
    - I + 1 - often a proline
    - I + 2 - often a glycine
  - Type II bond
    - I+ 1 - often glycine
    - I + 2
  - H bond between I and I + 3

carbonyl H-bonds to NH of i+3 residue

Reverse turns in polypeptide chains.

β turns

often glycine

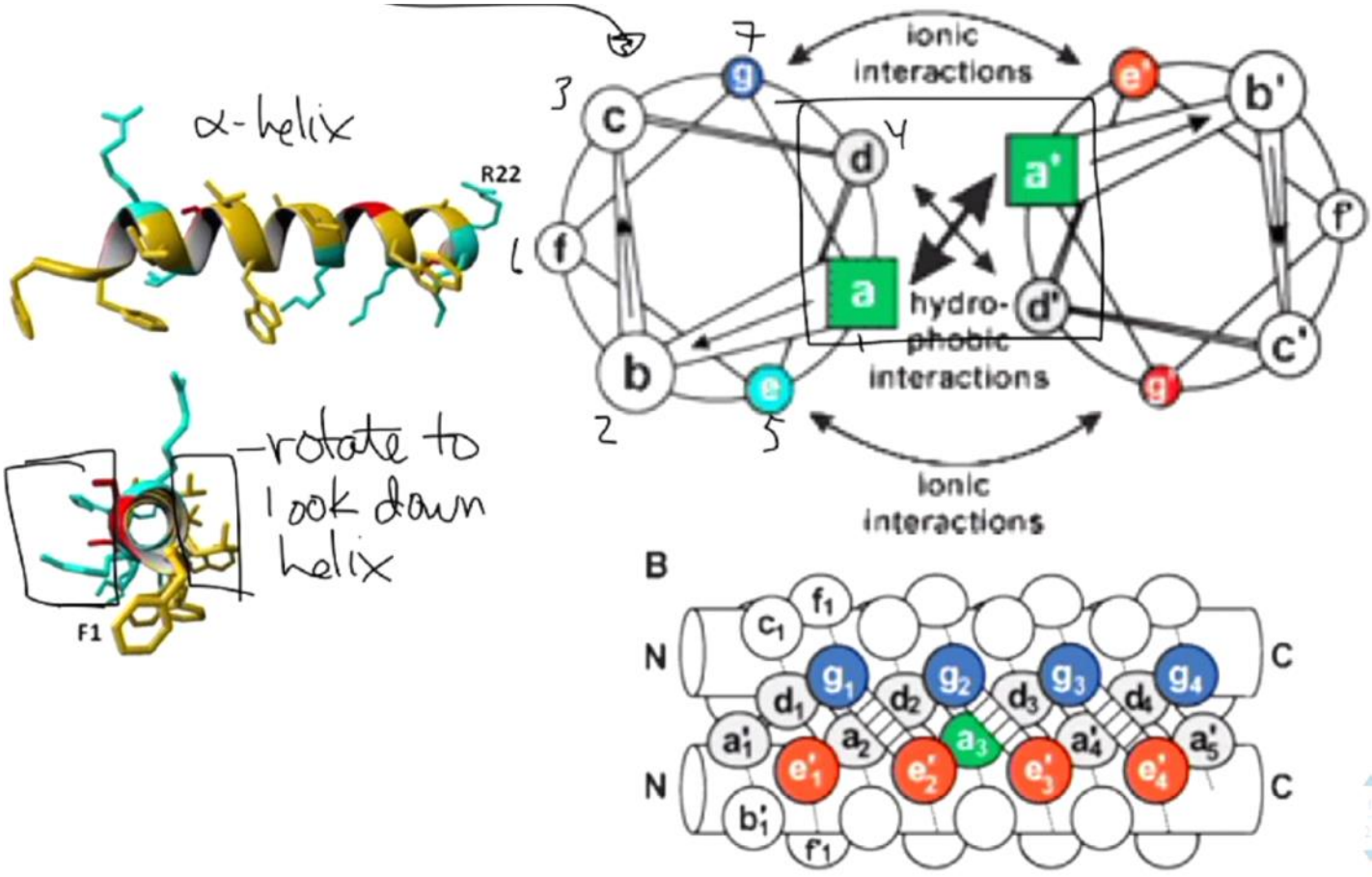


- Omega loops
  - Loop are 6 - 16 residues
  - Called omega loops
    - Structures that are not alpha helix, beta sheet or beta turn
  - On the protein surface - connect alpha helixes and beta strands
  - Involved in recognition
    - Proteins and biomolecular molecules bind to these
    - Unique structures
  - Can be flexible - move around
- Prediction of secondary structures
  - THEY CAN BE PREDICTED! - using a primary sequence
  - Based on empirical observation

Lecture 8 - Protein Folding

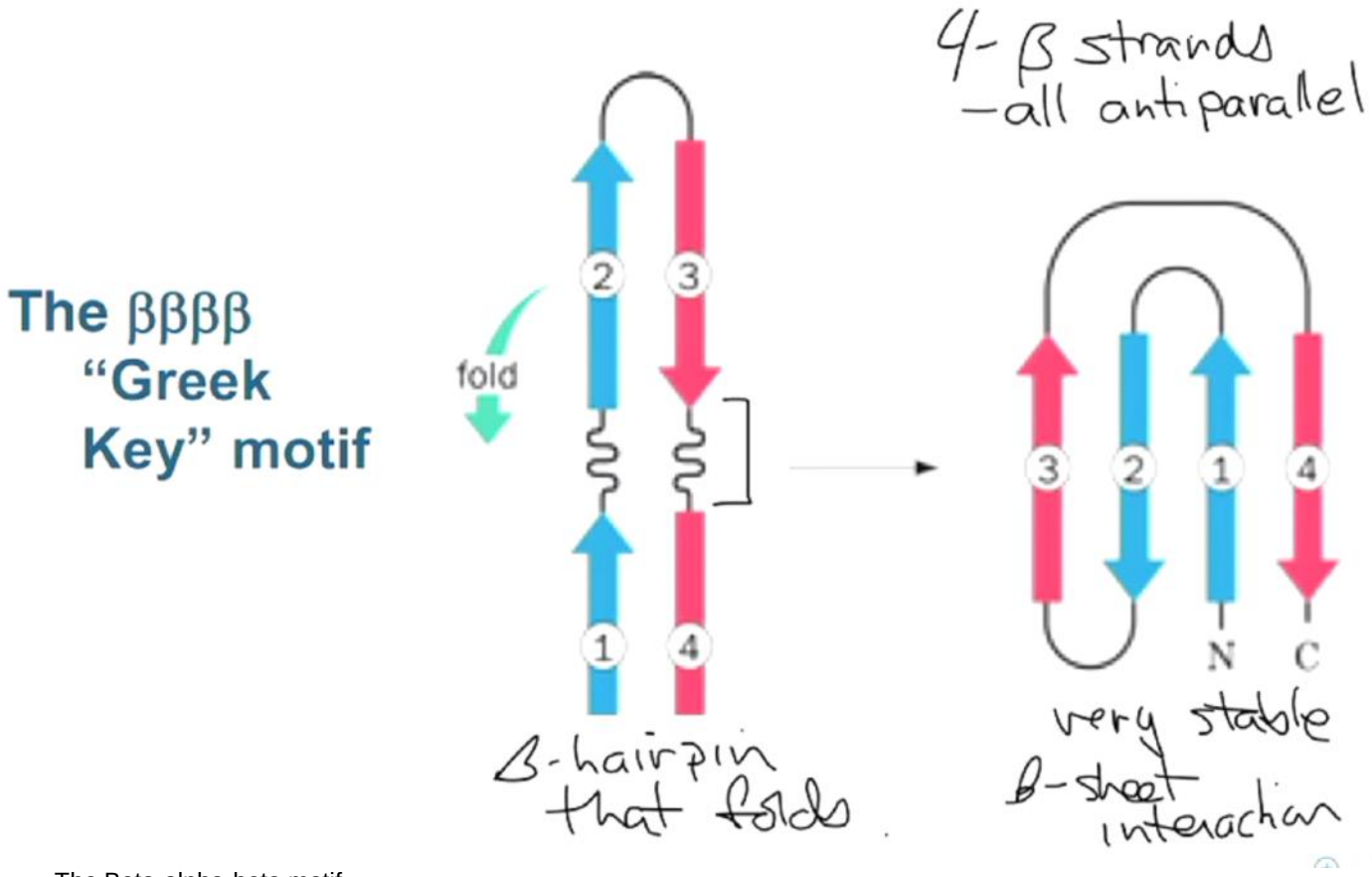
February 7, 2017 1:04 PM

- Combination of secondary structure elements
  - Motifs - supersecondary structures
  - Alpha - helices
  - Protein motifs
  - Beta strands - to make beta sheets
  - Loops and turns
- Are often repeated in same protein, different proteins
  - Tertiary structure - interactions between secondary structures
  - Only about 100 motifs
  - 4 very common - alpha- alpha, beta hairpin, greek key, beta-alpha-beta
- Certain motifs have associated biological functions
    - Eg. Helix-loop-helix
  - DNA binding or Ca2+ binding
- Secondary structures often depicted as ribbon diagrams
  - Helixes shown as ribbons
  - Strands - arrows
  - No cartoon depiction for turns or loops
- Common structural motifs of folded proteins
  - Alpha-alpha motif
  - Alpha-alpha hairpin
    - due to how sharp the turn is (2 - 3 residues)
  - Helix-loop-helix
    - Large turns - more resides in between
  - Helix-turn-helix
    - 12 or more resides
    - Helixes are at 90º
    - DNA BINDING - BIOLOGICAL FUNCTION or Ca2+ binding
  - Large binding surface between alpha-helixes
    - Usually hydrophobic
  - Stabilizes the motif - hydrophobic effect and london forces
- Coiled - coil
  - Two helices together - helical interactions (not necessarily alpha helix)
  - GON5 - histone aryltransferease
    - Regulates epigenetics
  - Helical interactions
    - DNA binding protein
    - Coiled-coil - the long interphase stabilizes interaction (ex. DNA binding)
    - Can predict interactions of one that are hydrophobic versus hydrophilic
- The helical wheel projection
  - Given an alpha-helix
    - We can compare the proteins looking down the helix
      - ☐ From the N terminus to the C terminus
      - ☐ A total of 7 residues in a wheel (2 full turns)
      - ☐ In the formation of alpha-alpha interactions
        - Hydrophobic core interactions (form a criss cross)
        - There is also ionic interactions on the periphery - formation of salt bridges
          - ◇ Strong salt bridges - less water due to hydrophobic core



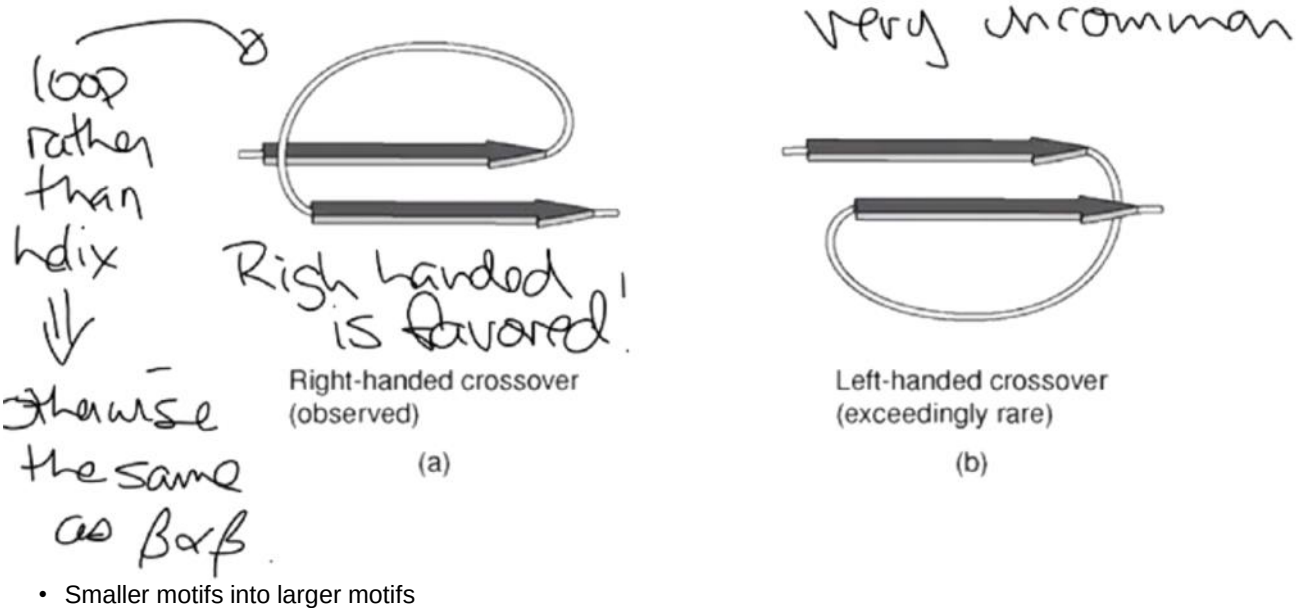
- ☐ Easy to design helixes that should interact in a helical wheel - formation of a coiled-coil

- The beta-beta motif
  - Bending at the I and I+3 residue of the hairpin
  - Antiparallel usually
  - Also called beta-hairpin
  - Strong linear H-bonds
- The Greek Key
  - Beta-Beta-Beta-Beta
  - The typical beta strand conformation cannot be formed causing a break, results in the formation of 4 beta sheets
    - Causes a break in the strand
    - 4 beta - strands are all antiparallel - thought to be the most thermodynamically stable form of any peptide sequence



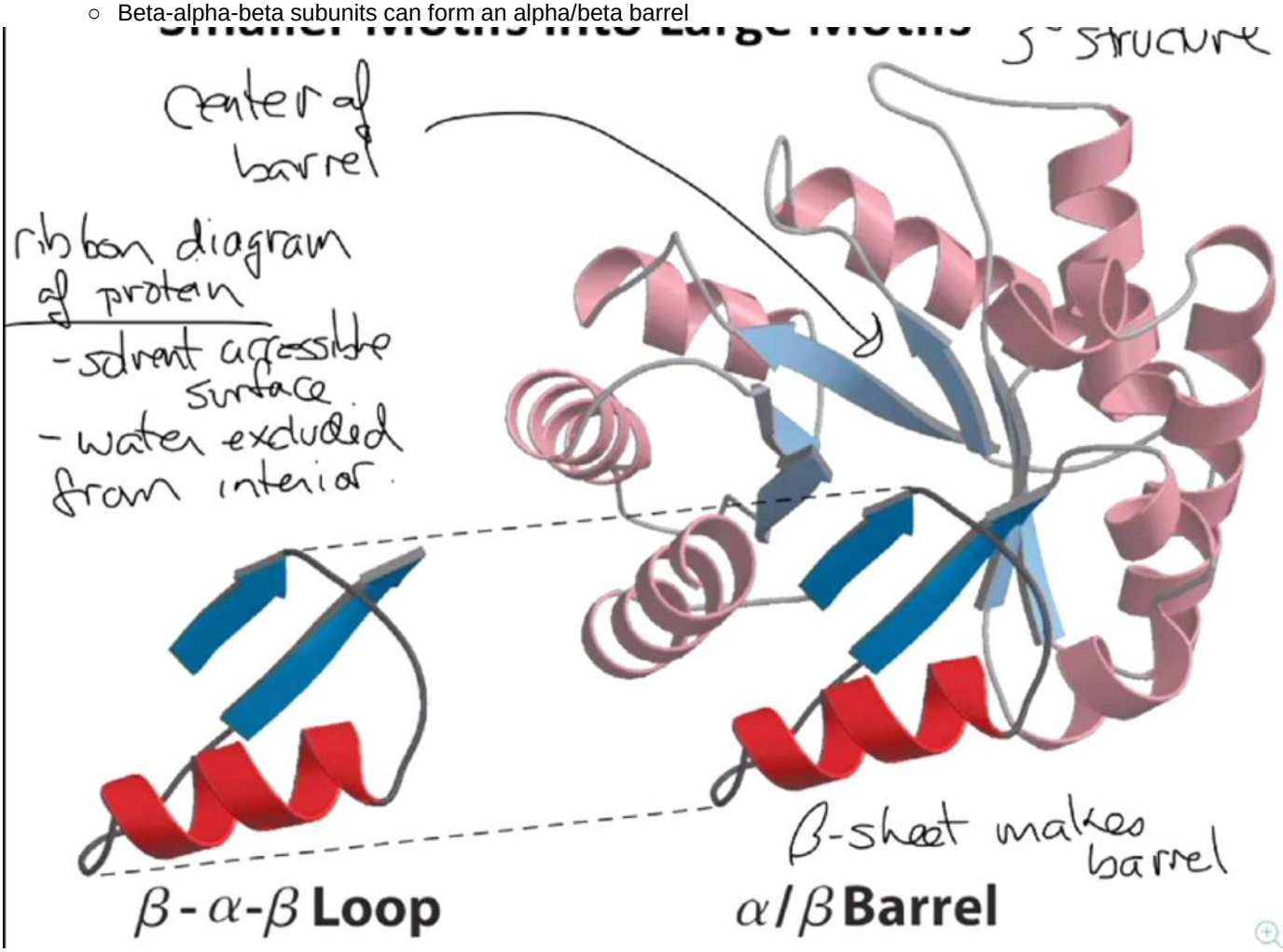
- The Beta-alpha-beta motif
  - Parallel B-sheet - tight turn between the beta sheet and alpha helix
  - Alpha helix connect both end of the strands together
  - Due to parallel structure - weaker H bonds
    - Stabilized by alpha helix which provides a hydrophobic binding site
    - Hydrophobic bases on BOTH the beta sheet and the alpha helix
    - HYDROPHOBIC BINDING SITE
    - Look at the helical wheel to identify hydrophobic sites and look at the beta sheet to identify hydrophobic sites - can form the Beta-alpha-beta motif
  - Usually a right hinded
- Parallel beta-sheets without a helix - beta loop beta motif
  - Loop rather than helix
  - Otherwise same as the beta-alpha-beta
  - Right handed is favoured and the left handed is very uncommon

Parallel β-sheets without a helix



- Smaller motifs into larger motifs

- Forms in a tertiary structure
- Beta-alpha-beta subunits can form an alpha/beta barrel



    - Alpha helices and beta sheets are on the inside of the protein - loops interact on the exterior
    - Water excluded from interior
- Protein families are characterized based on their motifs (all alpha ex. Serum albumin)
    - Look at insulin, HIV protease, Beta-amyloid, 3blm
- All beta - immunoglobulin fold - Homo sapien
- Alpha/beta - phosphofructokinase E. Coli - use alpha helix to connect beta sheets
    - Parallel beta strands
- Use alpha-helixes to connect parallel B-sheets (beta-alpha-beta)
- B barrel excludes water from fluorophore in GFP making it flourescent

- Quatenary Structure
    - Quaternary structure describes the organization of subunits in a protein with multiple polypeptides
        - Homo-multimers or hetero-multimers
    - Strands in quartnenary structures are labelled in greek letters, a subscript is used to indicate the number of monomers of that type
    - Subunits held together by non-covalent interactions
        - Oligomeric protein is more stable than dissociated subunits
        - Quatenary structures is the driving force for stability
    - Active sites often made up of AA residues from different subunits (at the junction of subunits) $\alpha\beta\gamma$
- Protein Structure can be dynamic
    - Structural changes can occur - big structural changes can occur as the protein functions
    - Adenylate Kinase
        - Substrate binding changes conformation of the protein
-

Lecture 9

February-09-17 12:37 PM

- Need to know secondary structures of pKa - ionic interactions
- Know the amino acid sequences
  - Protein Folding
  - A 3D shape dictated by the initial peptide sequence
    - FOLDING ALSO CONTROLED BY PRIMARY SEQUENCE
    - Unfolded random coil - to hydrophobic collapse (very fast) - molten globule (satisfying hydrogen bonding interaction by removing water from the interior, similar to oil) - side chains highly order - the final native fold
    - Hydrophobic collapse - very fast to decrease hydrophobic interaction with water of hydrophobic regions
      - Push hydrophobic regions inside
    - Molten globule is formed - sides chains are moveable - large volume
      - Hydrophobic regions are kept away from the water, hydrophilic interact preferentially
    - Eventually sides chains reach a highly ordered state to create the native fold
      - Small delta G - approx. -50 KJ/mol - only about 2 hydrogen bond difference
      - Small delta G allows the protein to achieve several conformations and dynamic states - biologically important
    - Denaturation
      - Proteins can be unfolded - denatured
      - Can be renatured but not always
      - INCLUDE SLIDE 1
- Folding is an extremely complex problem
  - How long would it take to fold if a protein had to sample all possible conformations?
  - Phi (φ) Psi (ψ) = 3 different conformations per amino acid
  - If a protein has 100 amino acids - 3^{100} conformations
    - $5x47^{10}$ conformations
  - Bonds rotate at 10^{-15} s
    - In a second can go through 10^{15} conformations
    - Try all conformations $5x10^{32}$ s
  - Lifetime of universe so far 4.3×10^{17}s - the amount of time taken to reach a stable conformation would take LONGER than the lifetime of the universe
  - Levinthal's paradox
- Cooperativity in folding
  - Reduces the total number of conformations that must be samples
  - After a interaction is formed (C1) - it causes I and J to be close to each other causing a possible non-covalent bonding interaction to occur
    - Causes formation of C2
    - This means that all other conformations are wiped out - none of the other conformations without C1 and C2 need to be sampled
    - Nucleation event of the initial structure templates of additional structure
- Example Coil-helix transition
  - α-helix formation
  - The very first turn is formed - the very first hydrogen bond is formed - nucleating the formation of the α-helix
    - Because it impacts the surrounding groups
  - When a hydrogen bonds forms - the H bond to the N becomes longer - greater partial negative charge in the N causing the carbonyl to become more partial positive and then causes the carbonyl to gain a greater partial negative charge
    - The initial hydrogen bond causes surrounding groups to be more reactive - to increase the need to H-bond due to a greater partial charge
    - Only need to nucleate a small portion of the structure to rapidly cause formation of the whole structure
  - Protein folding is a 2 state model

        - Due to cooperativity, there is no intermediate
        - Cooperative 2 state model - when it unfolds, the entire structure unfolds, if a region folds, the entire structure folds
    - Folding is initiated via hydrophobic collapsed
- Heat denatures proteins
    - ΔG must be positive
    - Take the ΔG change from unfolded to folded and change the negative sign to a positive sign
    - Heat is very important - vibrational energy
        - Breaks/weakens H-bonds
        - Decrease the strength of H bonds through bends
        - Sidechains begins to move - protein expands - London forces decrease and electrostatic interactions
    - Only a little has to be changed to disrupt the entire structure
    - Entropy is increased as well but it is mainly that ΔH becomes less negative due to weaker H-bonds, VDW and electrostatic
- Cold can also denature proteins
    - Hydrophobic effect - near freezing T, entropy of H_2O around non-polar residues is less different from those around polar residues
    - Hydrophobic effect decreases at lower T
        - H bonds are less dynamic - can decreases hydrophobic effects by 1/2 (less negative delta H) - delta G become positive
    - Organic solvents can also cause unfolding and ionic chemicals
- Energy surface for folding
    - Folding funnel - unfolded at high E
        - Protein folding occurs on a complex energy surface
        - Conformational change to the low E well
        - In the first diagram, must sample many states to find low E states
            - No Cooperativity, random sampling to eventually find the lowest energy structure
            - No driving force to find the lower energy structure
        - In the second diagram
            - Many, many paths follow to the downhill paths - increase in stability to achieve the lowest energy structure
            - Complex energy landscape - many downhill paths
            - Cooperativity - many pathways to low E
- Folding can go terribly wrong - due to many many intermolecular structure
    - Aggregated structures is the most stable NATIVE FOLD IS NOT THE MOST STABLE
    - Under unusual conditions - (very high protein concentration) can causes the formation of aggregation which is much more stable
        - But not a useful fold - the aggregates lose their function
        - Very stable and insoluble - becoming precipitates
        - Antiparellel β sheet - predominate the aggregate form
        - Formation of β amyloid plaques - several subunits aggregate to form these structures in Alzheimer's Disease
        -

Lecture 10

February 14, 2017 1:01 PM

- Protein purification
- Primary structure determination

- Electrophoresis: Analytical method for evaluating purity
  - Separation approach
  - #1 biochemical tool to analyse proteins
  - Analytical - cannot be used to purify proteins
  - Electric field which exerts an electrical force on charge molecules (they move in the +ve direction because prepared protein samples are NEGATIVELY charged)
  - More charge = fast migration
  - Force of friction counters the electric force
    - The protein must move through a matrix which will slow down its movement
    - Size of the protein is a factor
      - ☐ Large - greater friction force greater retardation
      - ☐ Smaller - smaller friction force lower retardation
  - THEREFORE
    - Very negatively charged and very small molecule will move very fast
  - DNA is also negatively charged - it can move through a gel
  - The gel
    - Made from cross linked polyacrylamide
    - PAGE
    - Cross-linking affects movement - level of cross-linking affects how fast protein move
    - Sits in a gel box - a buffer is used to movement a channel between the electrodes
    - The buffer
      - ☐ pH 9 - ensures peptides have negative charge
      - ☐ Current is applied
  - DNA gels
    - Agarose, pH 8 - phosphate backbone already has negative charge
- Protein are unfolded for electrophoretic analysis
  - SDS unfolds proteins (sodium dodecyl sulfate)
    - The protein is run under denaturing conditions
  - SDS wipes out the hydrophobic effect
    - Coats the hydrophobic residues
    - Solvates the hydrophobic residues
    - Destabilizing core of the protein fold - causes denaturing
  - Size is more uniform when unfolded
  - Adds extra negative charge - 1 SDS/2 Amino acids
  - Charge is more uniform and related to number of amino acids
  - SDS places a crucial role in ensuring that movement is due to the number of amino acids NOT the shape or charge
  - Reductant is added to convert disulfide bonds to free thiols - helps to unfold the protein
  - Native-PAGE - allows to determine quaternary structures
- Logarithmic relationship between molecular mass of a protein and its relative electrophoretic mobility in an SDS -PAGE
  - SPEED OF MIGRATION IS RELATION TO # of AMINO ACIDS
- Protein purification
  - Ion exchange chromatography
  - Gel filtration chromatography
  - Affinity chromatography
- Protein analysis by Digestion
  - To determine protein identity protein are digested into smaller fragments

    - Small peptide fragments can be easily analysed and sequenced (MS or Edman degradation)
- Trypsin Fragmentation
    - Most commonly used protease
    - Breaks amide bonds
    - Reacts at carbonyls of Lys(K) and Arg ®
        - At +ve charged residues
    - Produces a carboxylate and an amine (ammonium)
    - EXCEPTION
        - Will not digest if next as is a proline
        - K + R are frequent in proteins
        - Going to N to C - if the Arg-Pro OR Lys-Pro IT WON'T CUT
- Problems
    - What is the most likely secondary structure?
      M L Q S M V S L L Q
        - Underlined - non-polar structures
        - Not necessarily an amphipathic beta strand model - hydrophobic and hydrophilic should be alternating
        - Use the helical wheel to determine a partition in hydrophobic and hydrophilic sites
    - What secondary structure or motif does this sequence likely make?
      N Y K T R A E W K F E G P T L N V R I E L K G
    - Fits the β-strand model
    - Decide what model it fits the BEST - T is a situational amino acid
    - WATCH THE ECHO360 FOR Deciding where to stop a secondary structure
    - MOST LIKELY β-hairpin

Midterm 2 REVIEW

February 14, 2017 1:01 PM

- 20 questions
  - MCQ
  - Fill in the blank
  - Short answer
  - Model kits okay

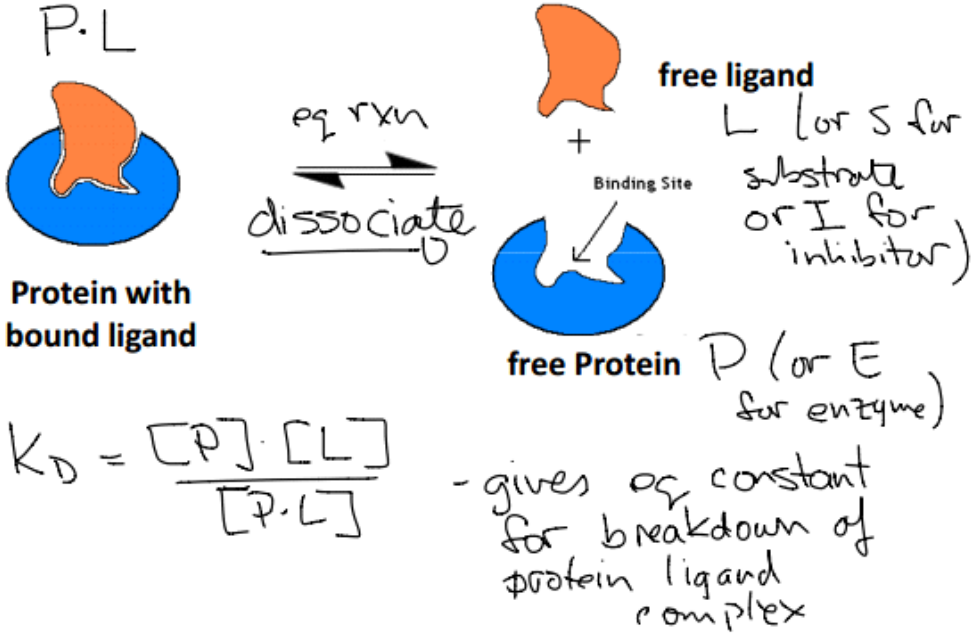
Lecture 11

February-28-17 7:42 PM

Protein Function

- Protein binding
- Enzymatic activity

- Molecular recognition, the ability of one molecule to "recognize" another through weak bonding interactions
  - All proteins bind to ligands
- Proteins can…
  - Bind to molecules for transport
    - Haemoglobin binds to O2 using a Heme group
  - Binds to substrate to catalyze a reaction
  - Bidns to ligands which modulate their activity
    - Potent inhibitor of HIV protease (SD146)
    - 14-H bonds inhibit the HIV protease (1BWB)
  - Binds to OTHER proteins
    - Antibodies bind antigens - targeting non-self such as infections
    - Quaternary structures are proteins binding to together
- General Descriptions of ligand binding
  - Simple equilibrium binding
  - Complex equilibrium binding
    - Cooperativity
    - Hill Plot
    - MWC and KNF model for cooperative binding
- General binding described by Dissociation constant K_D



- K_D would be given by the free ligand [L] and free proteins [P] over the initial concentration of ligand-protein complexes [PL]
- Why do we consider the dissociation constant of the ligand and protein over the association of the ligand and protein?
  - The units for K_D are mol/L

$$K_D = \frac{[P]\cdot[L]}{[PL]} = \frac{M \cdot \cancel{M}}{\cancel{M}} = M \quad \text{mol/L}$$

- $L + P \rightleftharpoons P{\cdot}L$

$$K_a = \frac{[P{\cdot}L]}{[P][L]} = 1/M \text{ units } L/mol$$

- less obvious what

  - K_a is in L/mol which is less intuitive than mol/L
- Similar to acid base chemistry (Handerson-Hasselbach)
  - When the [ligand] = K_D, 50% of protein is bound to ligand
  - Therefore [P] = [PL] so [P]/[PL] = 1
  - Similar to pka and pH where [HA] = $[A^-]$ then pH = pka

$K_D = \frac{[P][L]}{[PL]}$ at 50% binding $[P] = [PL]$

$\frac{[P]}{[PL]} = 1 \therefore K_D = [L]$

- like pka : pH $HA \rightleftharpoons A^{\ominus} + H^{\oplus}$ when $[HA] = [A^-]$ $pH = pka$

- Almost all binding sites are saturated if the ligand concentration is 10x K_D
  - 90% protein bound to ligand at this value

$K_D = \frac{[P][L]}{[PL]}$ $[L] = 10 \times K_D$

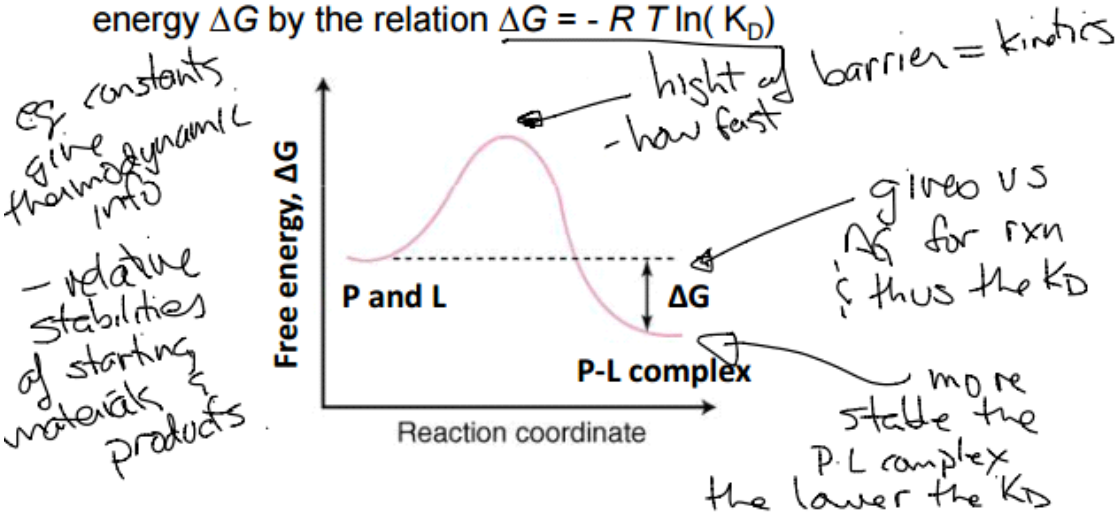
$\cancel{K_D} = \frac{[P]}{[PL]} 10 \times \cancel{K_D}$ $[PL] = 10[P]$

most protein is bound to ligand at $[L] = 10K_D$

- What happens when $[L] = 1/10\,K_D$ or $100K_D$

- The dissociation constant is relation to Gibbs free energy by the relation:
  - Starting position and final position of the reaction in free energy
  - Relative stabilities - equiliribum constants can tell us thermodynamic properties
  - The free energy of the PL complex is numerically related to the dissociation constant
  - The more stable the PL complex, the lower the K_D
    - If the binding site is destabilized, K_D increases
    - Greater interactions means less of the ligand is needed to bind to the protein
    - The smaller the K_D the tighter the bond
  - The height of the barrier tells us only the KINETICS

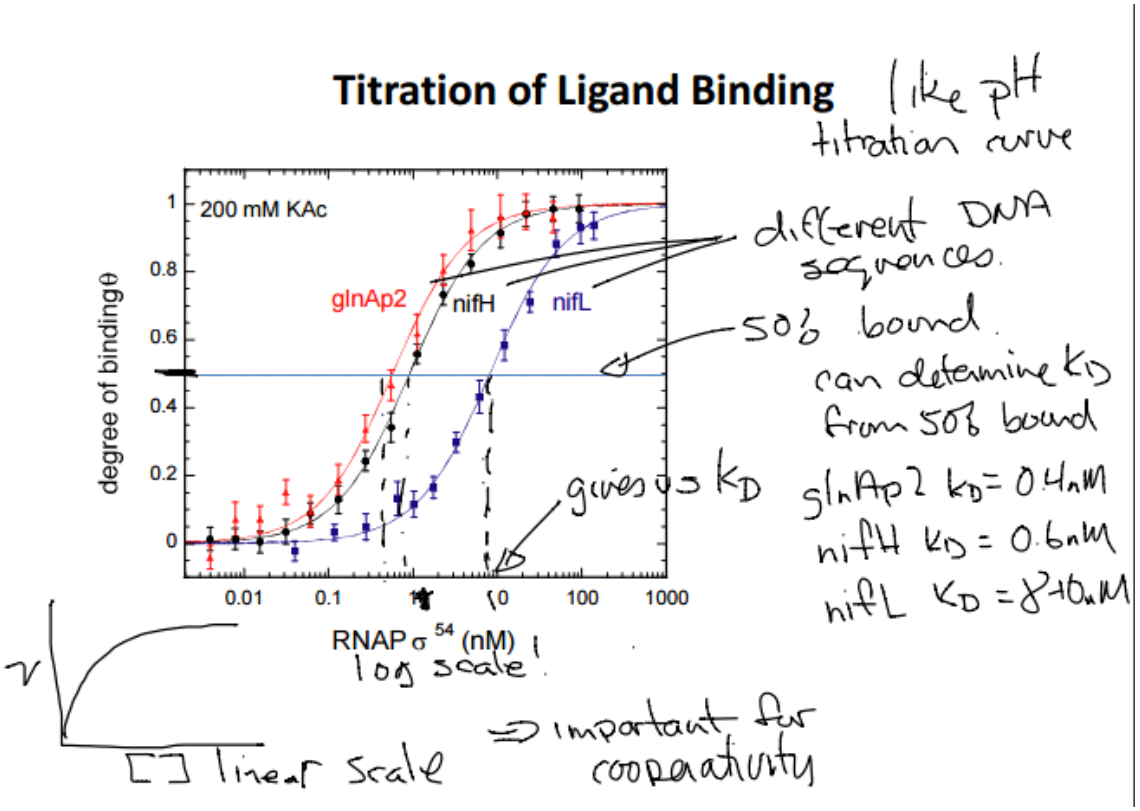
4. The dissociation constant K_D is related to Gibbs free energy ΔG by the relation $\Delta G = - R\,T \ln(K_D)$



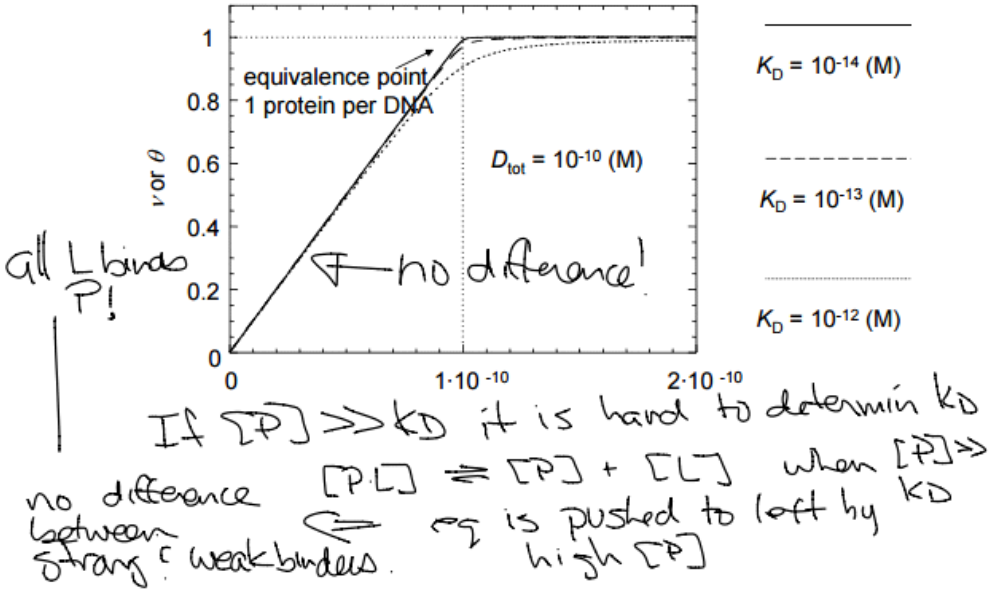
- K_D of a protein is often close to the cellular [L] that it targets
- Antibody-antigen interaction is the tightest bond
- Ligand Binding Curves
  - Degree of binding

$$\text{Degree of binding } (v) = \frac{[P\text{-}L]}{[P]+[P\text{-}L]}$$

  - The percent of the total protein bound to ligand (ligand-protein complex / (free protein + ligand protein complex))
  - To determine degree of binding
    - Using a fixed amount of protein so that [P] + [PL] = const.
    - Add a ligand and measure amount of [PL]
    - Determine degree of binding using equation above
- Titration of Ligand binding
  - The mid point gives us the K_D
  - The graph is in a log scale on the x-axis but can also be considered in a linear scale
  - In a linear scale it has a plateau at some point - WATCH OUT FOR LINEAR OR LOG SCALE!!!
  - You can use either the protein or the ligand at the x-axis because both are stabilized by the same interaction therefore would give the same dissociation constant

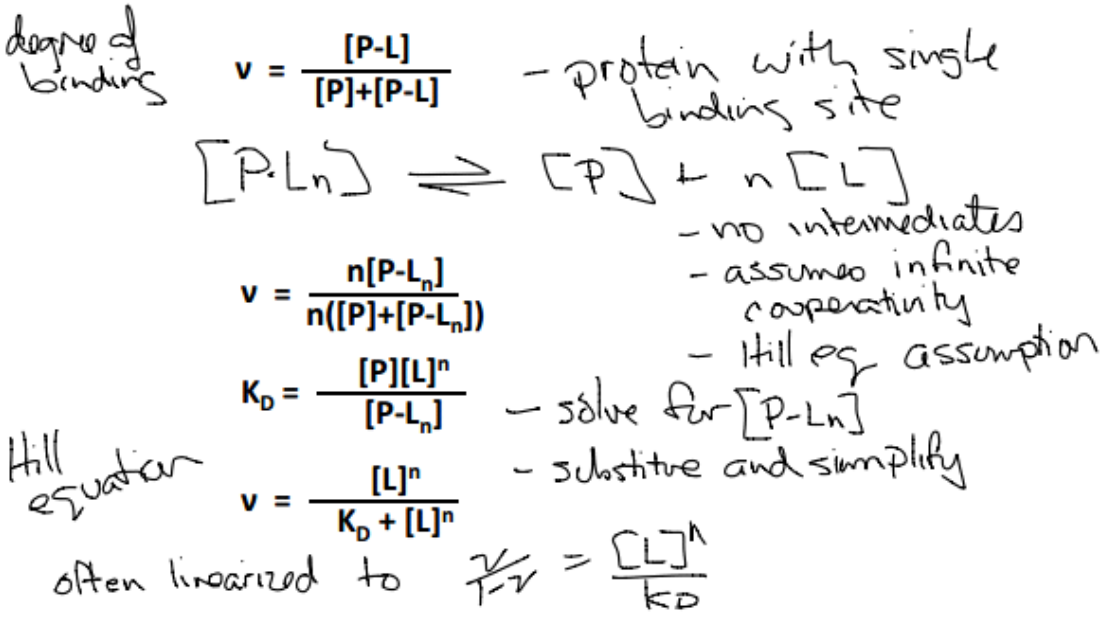


- If [P] >> K_D it is hard to determine the K_D
  - It is no longer an equilibrium rather a stoichiometric binding 1:1 ratio is reached as all L binds to P



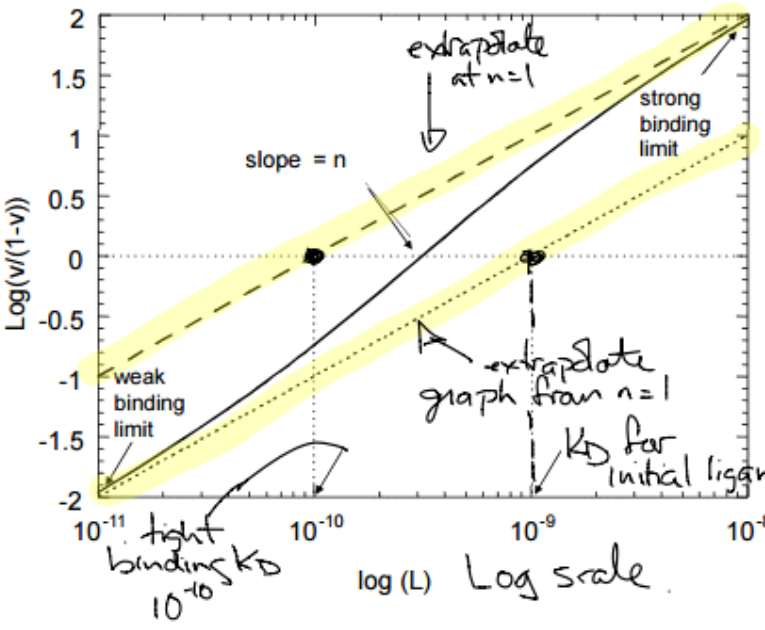
- Many protein have multiple binding sites
  - Multiple factors can be involved, protein binding sites, substrate and allosteric binding sites could all be involved
  - Nearly all protein have multiple binding sites - rarely the same - most are different
  - Simple K_D is when all binding sites are equivalent and independent
  - Co-operativity
    - All binding sites are equivalent and not independent
    - When one ligand binds to one binding site, it affects the other binding sites
      - Usually found in dimer, trimer….
  - Heterogeneity
    - Proteins that can bind several substrates at different sites with differing tightness- different binding sites that are non-equivalent
    - Heterogeneity binding sites when there are two different binding sites that non-equivalent
    - Independent because one molecule binding doesn't affect the other binding

site
        - Independent but not equivalent
    - When they both come together - all binding sites are non-equivalent and not independent
- Cooperativity
    - Binding of one ligand impact the binding of other ligands
    - Protein with n identical binding sites
    - [P - L_n]
    - Assumed that there is no intermediate - either bound to all n-Ligands or no ligands
        - Infinite cooperativity, Hill equation assumption



-

- Mathematical model to deal with n binding sites, when one binds all the other sites immediately binds, therefore no intermediate
    - Either fully bound or unbound
- Hill plot for cooperativity
    - The solid line is not straight due to cooperativity
    - Slope shows the cooperativity
    - Given in a log scale
    - The extrapolations are slopes on the curve which are equal to 1
        - Therefore these are concentration at which there is no-cooperativity
        - Therefore initially there is no cooperativity which can be used to determine the initial K_D for ligand binding = 10^{-9}
        - Second extrapolation in the strong bonding limit has a final K_D is 10^{-10}
        - The K_D lower ,therefore the binding is tighter the more molecules are bound to the protein
        - Extrapolate to log(v/(1-v)) to equal zero



slope of line = n

n = 1 non cooperative

n > 1 cooperativity

n < 1 negativity cooperative

Lecture 12

March-02-17 11:32 AM

- Hill equation
- Assumes infinite cooperativity
  - Binding of the first ligand causes immediate binding of all the other ligands
  - No intermediate, either fully bound or unbound
- n in the hill equation is a measure of cooperativity (NOT THE NUMBER OF LIGANDS)
  - Doesn't tell you how many ligands are bound
  - If it was truly infinitely cooperative it would tell you the number of ligands bound
    - Example infinitely cooperative Haemoglobin is but is actually 2.8

Protein with n ligand binding sites

$$v = \frac{[L]^n}{K_D + [L]^n}$$ Hill eqn

$P L_n \rightleftharpoons P + nL$

- n = number of ligands (greater than 1)
- assumes infinite cooperativity

n = measure of cooperativity

- linearized

$$\log(v/(1-v)) = n\log[L] - \log K_D$$

$y = mx + b$

y intercept $= -\log K_D$

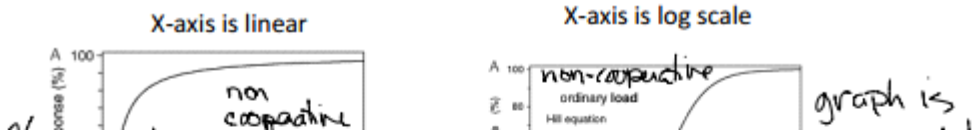
$m = n$

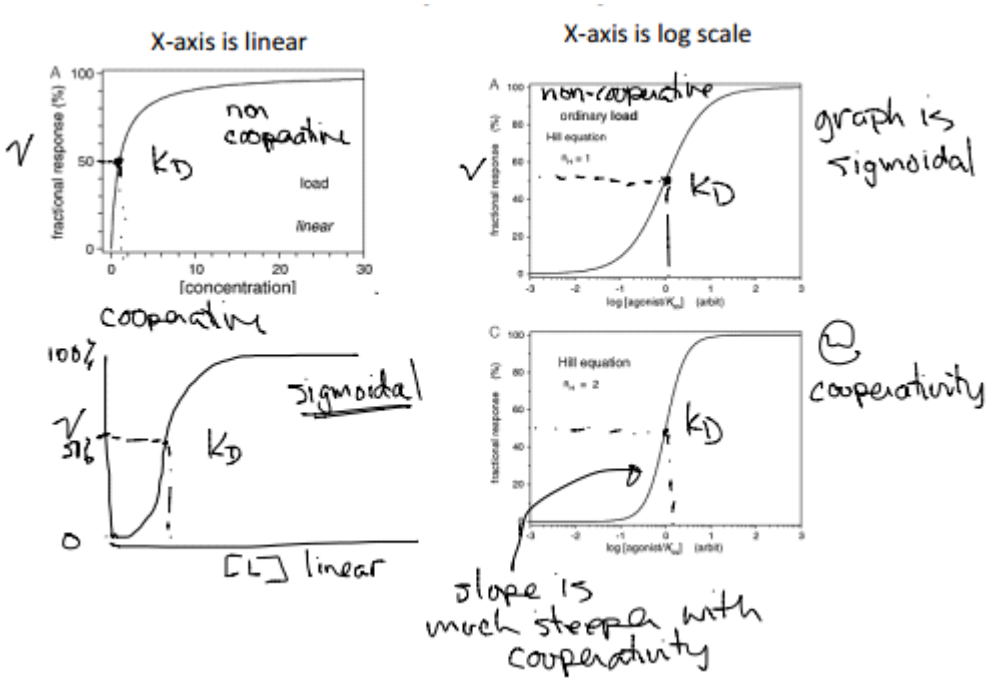
n = 1 non cooperative

n > cooperative

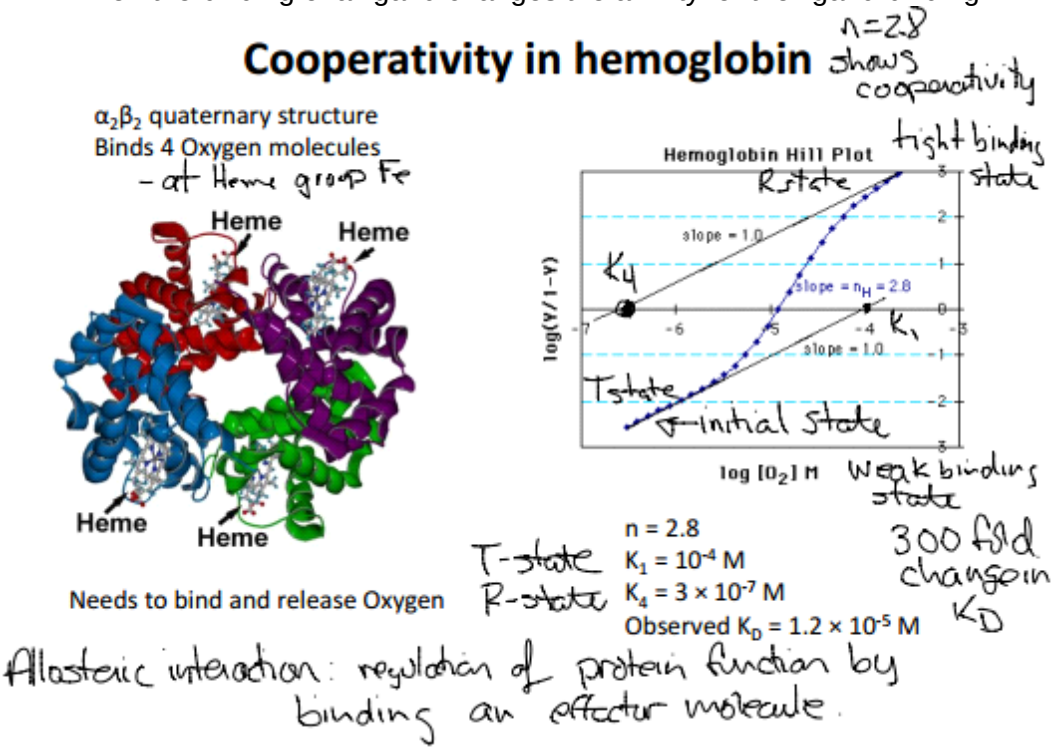
n < negative cooperativity

- The Hill plot cooperativity
  - The slope changes in the data plotted- not a straight line
  - Cooperativity changes as ligand concentration changes
  - At the low ligand concentration and high ligand concentration the slope is equal to 1
    - At these spots, ligand binding is in a non-cooperative fashion
    - K_D can be found from the extrapolates of the low and high concentration states
    - v=0.5, so log(v/(1-v))=log(1)=0
    - Where the extrapolate crosses the x axis is the K_D
  - In the middle the slope is greater - there is a cooperative ligand binding
  - The K_D has changed, weak binding at low concentration and strong binding at higher ligand concentration
    - Cooperativity effect changes the way the ligand binds
  - There are 2 states
    - Weak binding state at low ligand concentration
    - Tight binding state at high ligand concentration
    - A cooperative effect in between to change the binding strength
- Hill equatiion: other ways to visualize cooperativity
  - V is always 50% at K_D
  - In a cooperative system:
    - V vs. axis
      - ☐ It followed a sigmoidal shape on the linear function
      - ☐ Can obtain an AVERAGE K_D at 50% protein bound
  - In a linear scale - it is easy to find what Is non-cooperative and what is cooperative but it is difficult to do so on a log scale



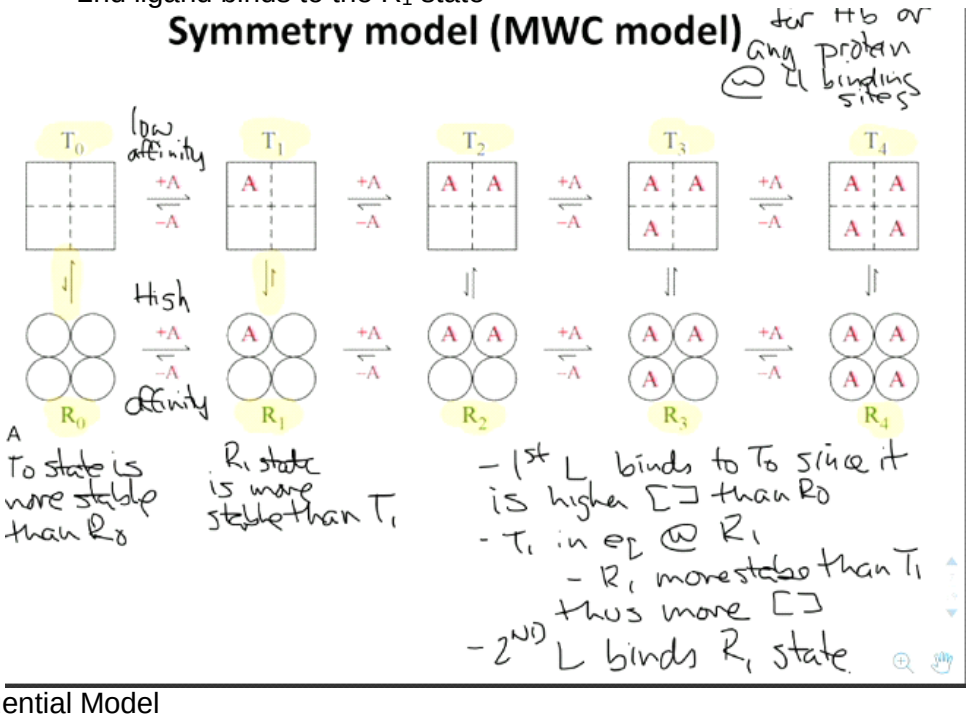


-
  - Cooperativity in hemoglobin
    - Slope of 2.8 which shows cooperativity
    - In the lung, there is a higher oxygen concentration so tight binding occurs (low K_D)
    - In tissue, there is low oxygen concentration causing high K_D and promoting dissociation
    - Higher state is called the R state and the lower state is called the T state
    - Allosteric interaction: regulation of protein function by binding an effector molecule
    - This form of interaction is called an allosteric interaction - cooperativity is when the binding of a ligand changes the affinity for the ligand binding



- Models to explain altered binding in R and T state
  - Symmetry model (MWC model)
    - General model to describe allosteric
  - Allosteric protein is symmetrical - assume all subunits are the same (in Hemoglobin for example)
    - Assume all 4 binding sites are exactly the same - treating them as 4 identical sites

- Protein can exist in 2 state which are in equilibrium
  - The states are in T and R
  - Based on the absence of ligand and when ligand is bound
- Ligand can bind both states
  - It can bind in both the T and R state
  - Binding can shift equilibrium between the T and R states

- The symmetry model states that all sites have to be the same, either all T state or R state
  - In the Hill-plot, haemoglobin the T state binds oxygen with low affinity, R state binds ligand with high affinity
  - At low ligand concentration, the T state is binding because we see low ligand binding state
    - So T_0 is the state that will be in greatest abundance, greatest likelihood to bind to the T state over the R state
  - Once ligand binds, it changesto T_1
    - Shifts the equilbrium to the R state, shown in the magnitude of the equilibrium arrows the R_1 state is more stable, therefore there is a greater amount of the R_1 state (greater concentration)
    - As more ligand binds, it causes more of the T state to shifts towards R state
    - Greater binding affinity as greater portion of the population is in R state
    - 2nd ligand binds to the R_1 state
  - 

- Sequential Model
  - General model of allosteric binding
  - Subunits of multimeric protein have 2 conformational states
    - Low affinity and high affinity state
  - Ligand binding causes a conformation change
    - STATES ARE NOT AT EQUILBRIUM AS IN SYMMETRY MODEL
    - The ligand causes the conformational change from the low to high affinity state
    - Due to induced fit binding
  - The conformational change in one subunit causes the other subunit to change into an intermediate state with higher affinity IF POSITIVE COOPERATIVITY
- Induced Fit versus Lock and Key
  - Lock and key mechanism - precise site that can bind a ligand and fit together like a lock and key
    - No change in conformation, static model of protein structure
    - Refuted because if a substrate can bind effectively then it could never reach a transition state
  - Induced fit
    - The conformation might not be ideal for binding the substrate but as they

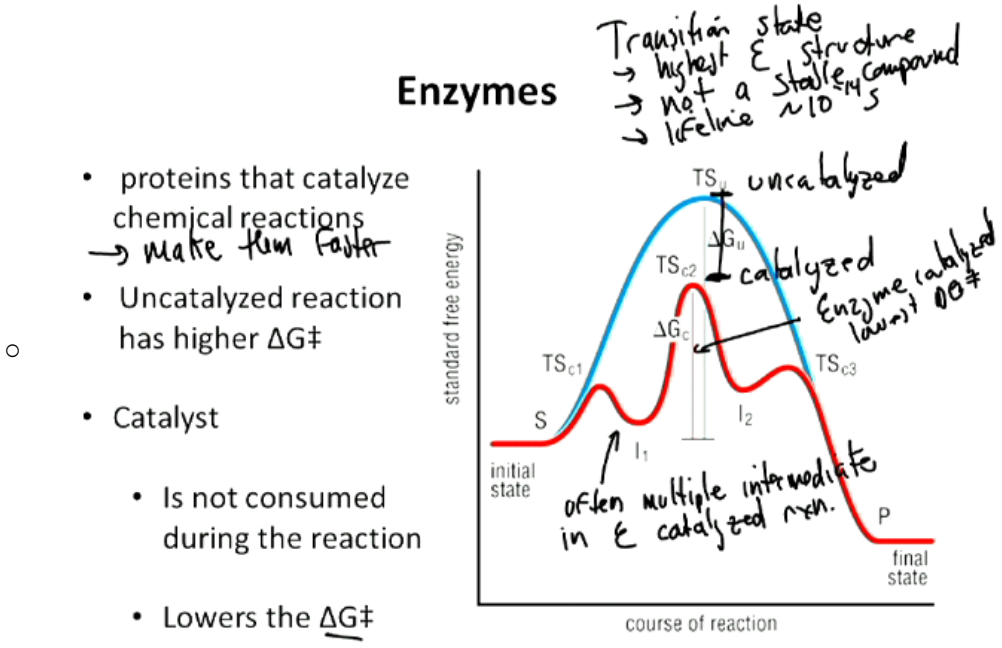
get closer, interactions cause a change in conformation allowing for substrate binding

- The first protein is in the low affinity T state
    - Binding of a ligand to the T states causes the neighboring subunits to change in conformation to a strained intermediate structure - greater affinity
    - Strained/destabilized states are more likely to bind with a ligand, changing to the high affinity states and causing other subunit to enter an intermediate state
    - Final T state subunit has similar affinity to R state and quickly binds the final ligand
    - The level of cooperativity depends on the strength of mechanical coupling of subunits
        - In the other model it has more to do with the equilibrium, change in stability between the T and R state
- BOTH SYMMETRY AND SEQUENTIAL MODEL ACCOUNT FOR HAEMOGLOBIN ACTIVITY

Lecture 13

March-07-17 11:41 AM

- Protein Function
  - Enzymatic activity
    - Michelis Meten Kintetics - Gibbs free energy graphs
  - Mechanisms of Catalysis
    - How enzymes enhance reaction rates
  - Inhibition
- Enzymes catalyze chemical reactions
  - Transition state in between
  - Speeds up the process to reach equilibrium by lowering the TS state energy
    - Lowers the activity energy
  - Transition state is the highest energy structure and it is not stable compound, lifetime is approx. 10^{-14} sec
  - Often multiple intermediates in the enzyme catalyzed reaction
    - Enzyme and substrate
    - Transition State
    - Product and substrate



- Catalyzed rate is massively increased from the uncatalyzed rate
- Some enzymes are so fast, delta G is almost zero
  - Everytime enzyme interact with the substrate product is immediately formed
  - Effectively like diffusion, no activation energy - no transition
  - Rate is limited by ONLY be how fast the enzyme and substrate interact
  - DIFFUSION LIMITED 10^8 - 10^9 M^{-1} s^{-1}
- Enzyme catalyzed versus uncatalyzed
  - Uncatalyzed as [S] increase the velocity increases
    - Rate = k[S]
  - In catalyzed, two phases, the linear phase where rate increase as [S] and a plateau, Vmax
  - Eventually reaches a plateau a Vmax
    - At low S conc. Lots of available enzymes that the S can bind
    - As concentration increase, reaches a saturation point of the enzymes, a point at which all the enzymes are bound with substrates
- Order of reactions
  - 1st order: rate= K_1[S]
  - 2nd order: rate = k[S][E]

- 3rd order: rate = $k[S]^3$
- Mechanistic model for enzyme catalysis
  - Enyzme first binds the substrate
    - E+ S <-> E * S
      - ☐ Like ligand binding,
  - Enzyme catalyzed formation of product
    - E * S <-> P + E
      - ☐ Like a chemical reaction
-

Mechanistic model for enzyme catalysis

2 steps

- enzyme bind substrate first

$$E + S \rightleftharpoons E \cdot S$$

. like ligand binding

- enzyme catalyzes formation of product

$$E \cdot S \longrightarrow P + E$$

. this step is like a chemical reaction.

→ Can have other intermediates:

$$E + S \rightleftharpoons E \cdot S \longrightarrow E \cdot P \rightleftharpoons E + P$$

* Useful for product inhibition!

-
- Michaelis-Menten Model
  - Rate constant that describes the formation of the ES complex
  - Rate constant that describe the formation of product k_{-2} <<<<<k_2
    - So it is ignored
  - $K_{eq}= k_{-1}/k_1$
    - Reversible binding
    - Rate of P formation = $k_2[ES]$ - difficult because [ES] is a TS structure

Mechanistic model for enzyme catalysis

Michaelis-Menten model — most common model for describing enzymes.

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} P + E \qquad k_{-2} \text{ cccc } k_2$$

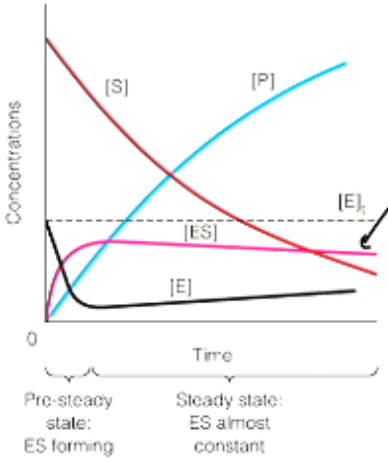
-
$$K_{eq} = \frac{k_{-1}}{k_1}$$ reversible binding

$$\text{rate of P formation} = \frac{dP}{dt} = k_2[ES]$$ * hard to measure $[E \cdot S]$

$$[E_{Total}] = [E] + [ES]$$

- Once the rxn starts [ES] is constant - this is the steady state approximation
- Formation [ES] = consumtion of [ES]
  - As product foms, substrate binds on to reform [ES]
  - Rate of consumption can be product foration or dissociation
  - Formation of[ES] is effectively equal to the CONSUMPTION of [ES]

Steady state assumption



$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} P + E$$

→ once rxn is going [ES] is constant

→ steady state approximation (SSA)

Formation [ES] = consumption [ES]

$$k_1[E][S] = k_{-1}[ES] + k_2[ES]$$

$$k_1[ES] = (k_{-1} + k_2)[ES]$$

→ solve for [ES]

Relate [E] and [ES] to $[E_{total}]$

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} P + E$$

→ still do not know free [E]!

$$[E_{Total}] = [E] + [ES]$$

- → from the SSA: $[E] = \frac{(k_{-1} + k_2)[ES]}{k_1[S]}$

So: $[E_T] = \frac{(k_{-1} + k_2)[ES]}{k_1[S]} + [E\cdot S]$

→ solve for [E·S]: $[E\cdot S] = \frac{[E_T]}{\left(1 + \frac{k_{-1} + k_2}{k_1[S]}\right)} \Rightarrow \frac{[E_T][S]}{\left([S] + \left(\frac{k_{-1} + k_2}{k_1}\right)\right)}$

Final Michaelis-Menten rate equation

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} P + E$$

→ since $\frac{dP}{dt} = k_2[ES]$

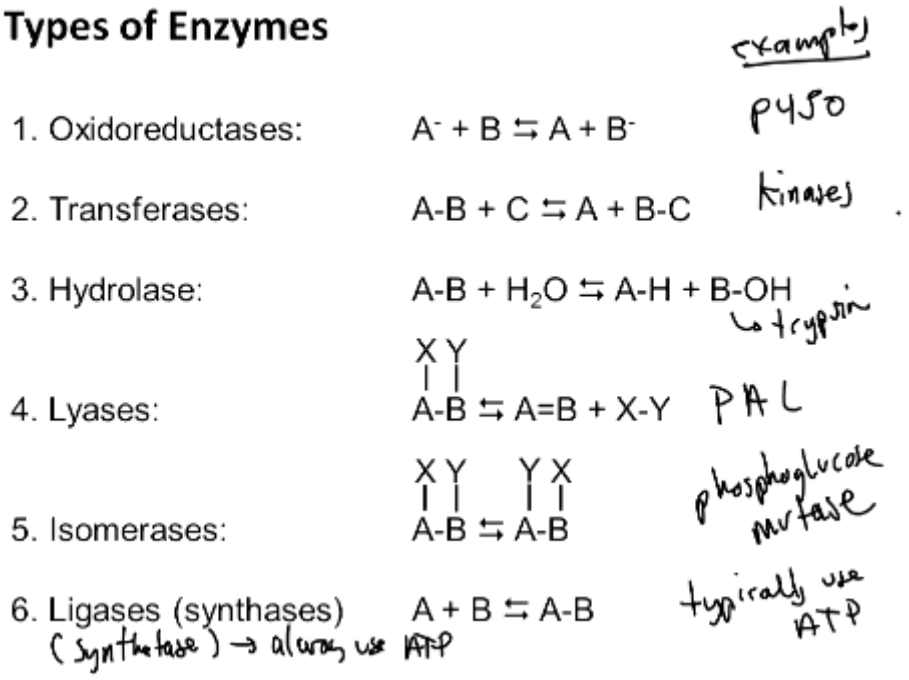
K_M = Michaelis constant $= \frac{k_{-1} + k_2}{k_1}$

- $= \frac{k_2[E_t][S]}{[S] + \left(\frac{k_{-1} + k_2}{k_1}\right)}$

$= \frac{k_2[E_t][S]}{[S] + K_M}$ $\quad V_{max} = k_{cat}[E_t]$

$$M.M.\ eq = \frac{V_{max}[S]}{[S] + K_M}$$

- K_{cat} is equal to the turnover the rate - it is also equal to k_2
- If $k_{-1} >> k_2$
    - The rate of dissociattion is much greater tha the rate of formtion of product
    - Therefore $k_{eq}=k_D$
        - k_D is the dissociation constant
- The difference in energy of the free enzyme and substrate is determined by the specific constant
    - K_{cat}/k_M = specificity const.
    - Gibbs energy of activation - can be determined by the specificity constant, the energy required to form the product from the substrate
    - Delta G (cat.) is from k_2 or k_{cat}
        - The energy required to convert the substrate to the products
        - K_3 tend to be very fast to dissociate the product and free enzyme
- Types of Enzymes
    - Oxidoreductases - oxidative reactions
    - Transferases - trasnfer a group from one molecule to another
    - Hydrolase - breaks bonds through hydrolysis
    - Lyases - molecular broken apart by putting in a double bond
    - Isomerases - changes the isomer
    - Ligases - creates bonds, synthases always use ATP

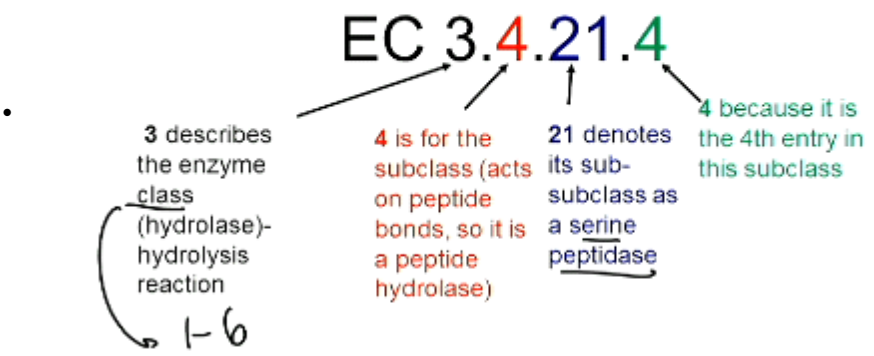


Enzymes

assigned an EC number which defines exactly which reaction it catalyzes

EC - enzyme commission

Eg trypsin is **EC 3.4.21.4**

- 

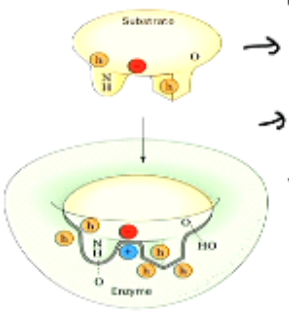
  - Enzyme Nomenclature Database - http://expasy.org/enzyme

- Co-factors
  - Non-protein part of enzyme
  - Ions: Zn2+
  - Co-enzymes - organic molecules - chemically changed by the enzyme
    - Co-enzyme A, riboflavin, lipolic acid
  - Prosthetic group: tightly or covalently bound co-factor (heme group)
    - Apoenzyme: inactive enzyme lacking its prosthetic group (heme without Fe2+)
    - Holoenzyme: apoenzyme plus its prosthetic group (active) (heme with Fe2+)

Enzymes are highly substrate selective

induced fit

- **Active site**: where the reaction occurs
- **Binding site:** where substrate binds
- **Catalytic site**: where reaction occurs



→ size & shape

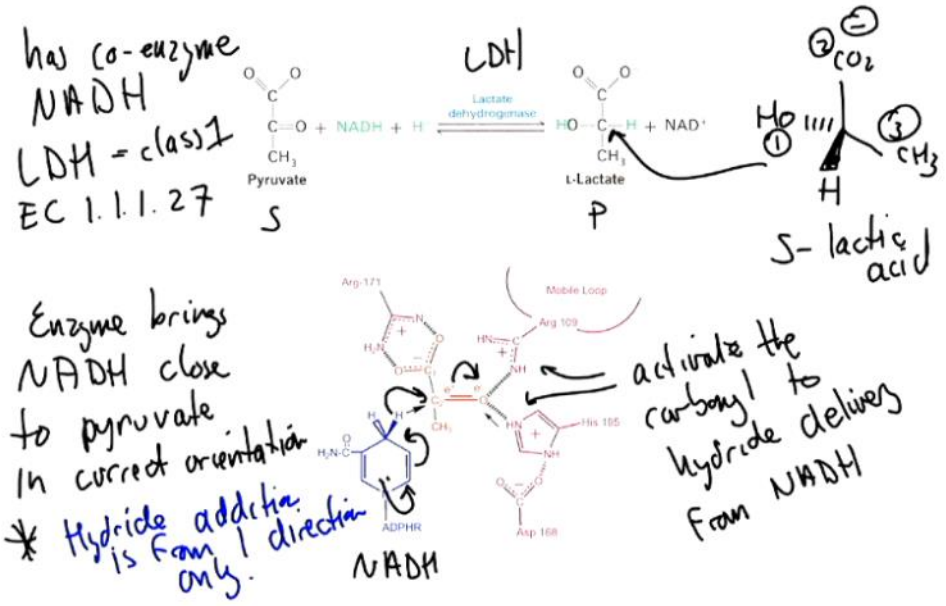
→ H-bond

→ electrostatic interaction

Binding site contacts substrate & non-covalent interactions

- Enzymes are highly stereoselective
  - So right size, shape and stereochemistry
  - Hydride addition can only occur in one direction as a result only the S isomer is produced

Enzymes are highly stereoselective

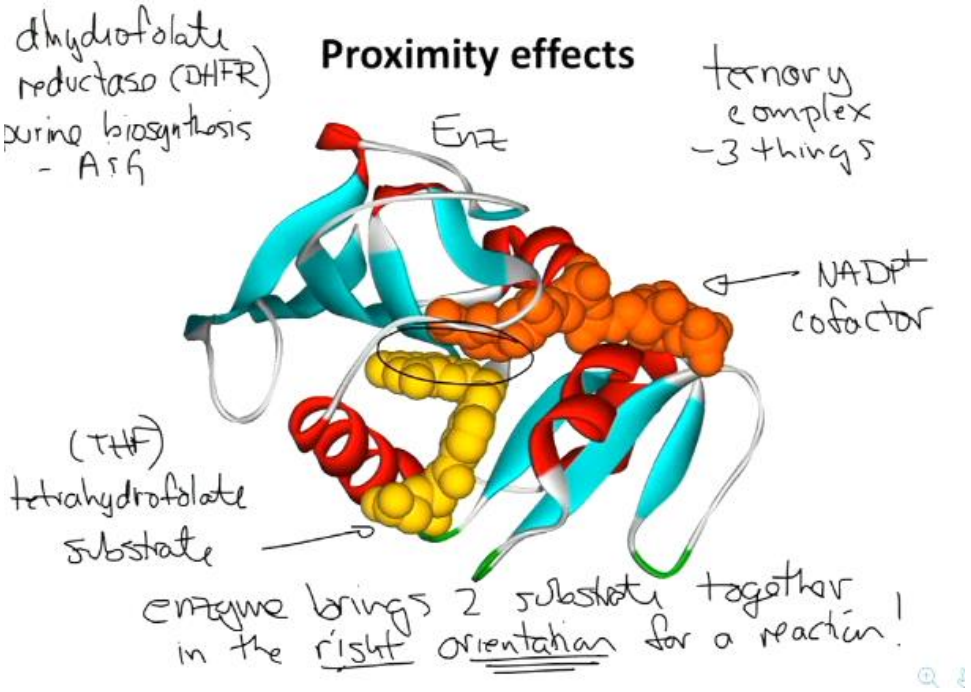


- 6 Key methods how an enzyme increases reaction rate
  - Proximity - bring everything closer
  - Acid-base catalysis- source of protons
  - Metal ion catalysis- positive charge coordinate two group
  - Covalent catalysis - transient covalent bond
  - Electrostatic catalysis - two opposing charges
  - Transition state stabilization - stabilize the transition state structure that is ideal

Lecture 14

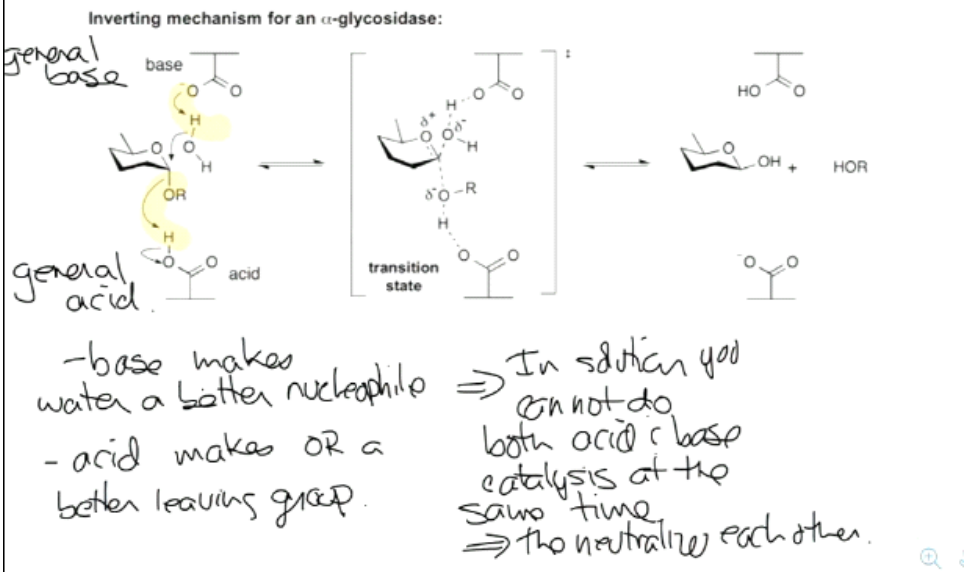
March-09-17 12:29 PM

- Enzymes
  - Michaelis-Menten parameters give relative energies of S+E, E-S and TS

- Proximity Effects
  - Brings molecules together and in the correct orientation
  - Dihydrofolate reductase - involved in A and G synthesis
  - Reactive sites are in very close proximity
  - Essentially increases the concentration of the substrate
    - Impact on the overall entropy of the reactions by favouring productive collisions from occuring
- 

- Acid-base catalysis
  - Can separate an acid from a base - can be spatially separated so they never fully react
    - Means that within the enzyme both base acceleration and acid acceleration can be used because there is no neutralization
    - Acid can be used to activate the electrophile, a nucleophile can be activated by a base
      - ☐ Dramatically improve reaction rate
- At the top part, the carboxylate acts as a base (nucleophile)
  - Bottom part acts as a acid and improves the reactivity of the leaving group (alkoxide group -OR)

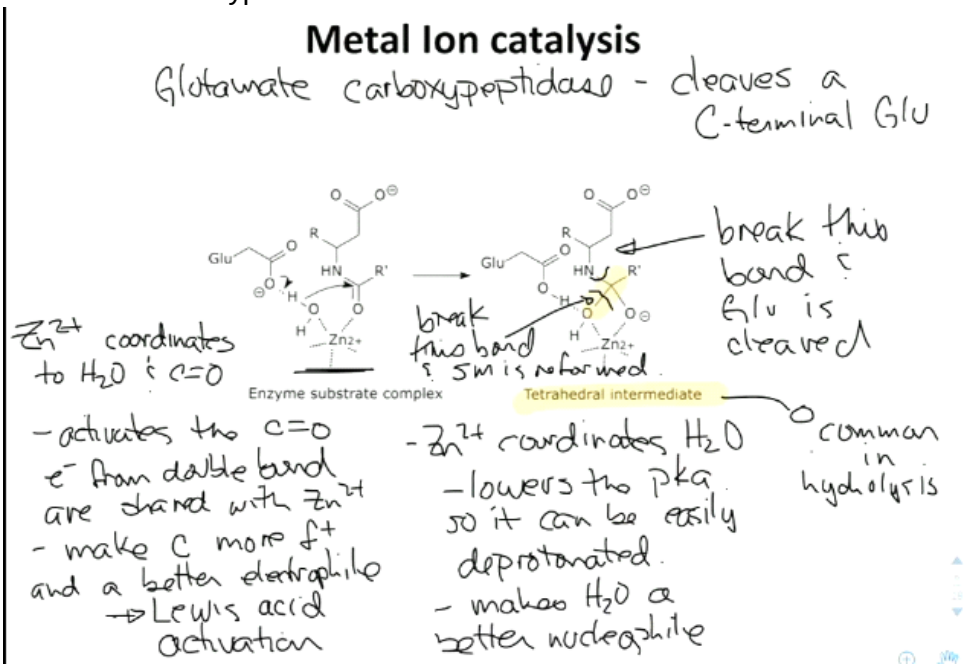
Acid-Base Catalysis



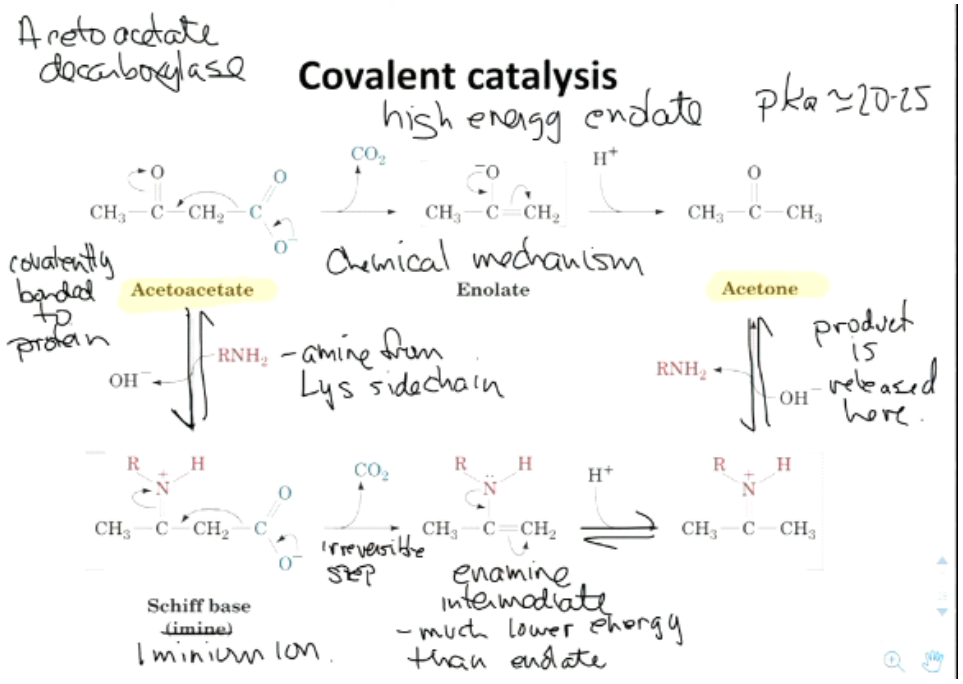
- Amino acid that do A/B catalysis are:
  - Asp, Glu, His, Cys, Tyr, Lys

- Metal Ion Catalysis
  - Lewis acid type activation

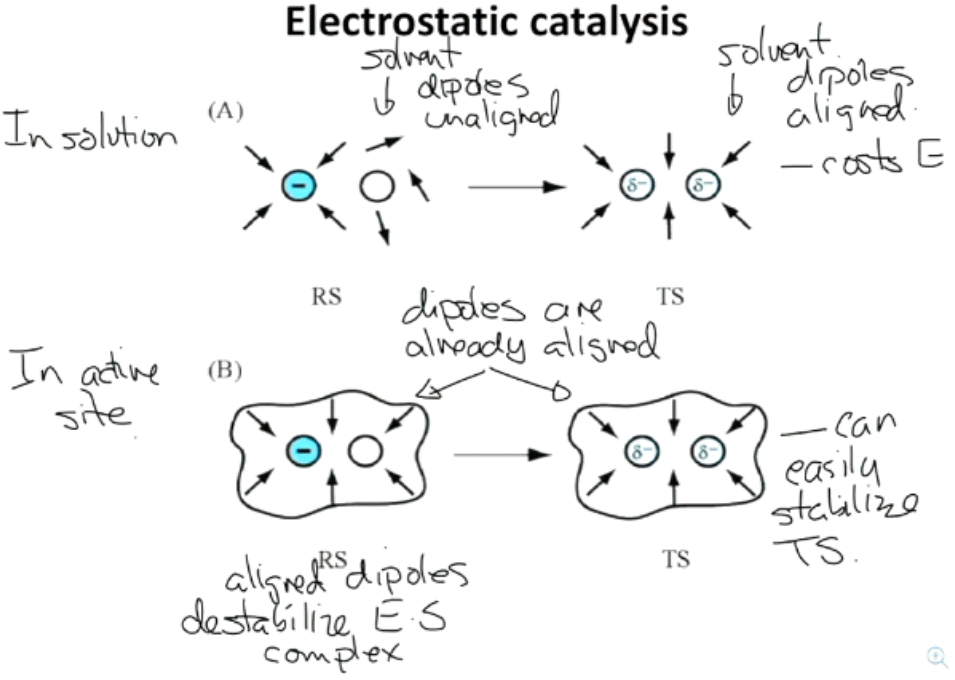
Metal Ion catalysis



- Covalent Catalysis
  - Goes from Acetoacetate to acetone
    - Through a chemical intermediate called enolate
    - A deprotonated keynote
      - ☐ pKa is approx. 20-25 - not possible to form an enolate in water - very high energy intermediate
  - Enzyme can create new covalent bonds to change the TS state structures
    - Amine from Lys - acts reversbility to form an imine

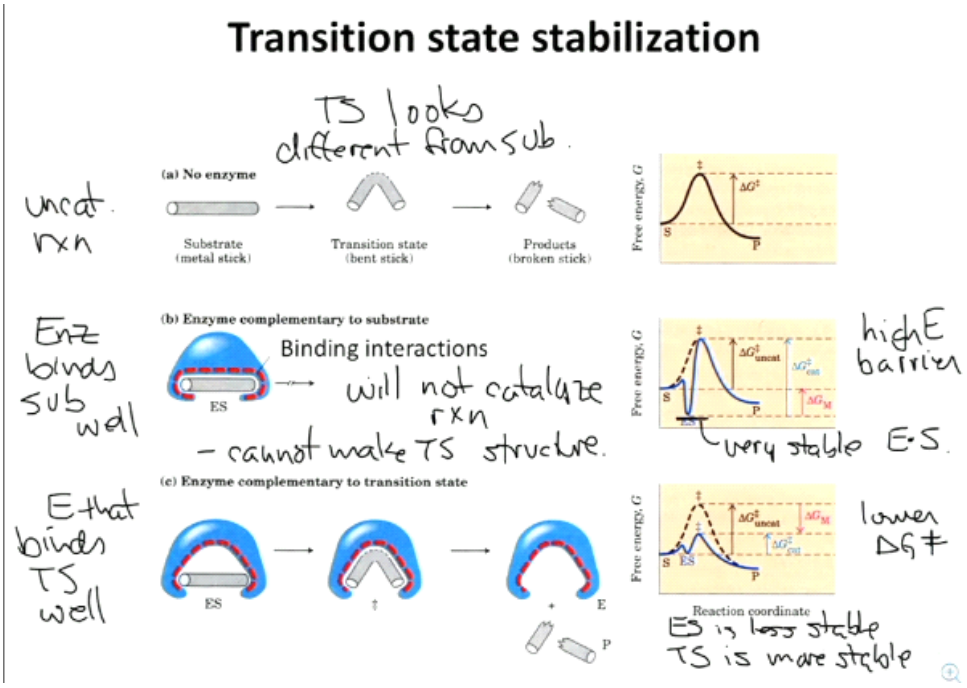


- Electrostatic catalysis
- In solvent, charged species are stabilized by aligning the dipoles of the solvent with the charged molecule
- In neutral molecule, randomly aligning
    - As the TS occurs, the solvent molecules have to realign to stabilize the TS, the distribution of charge across the two reacting species
    - This takes up energy and slows down the reaction - energetic penalty to realign the solvent molecules
- In an active site: everything is held together by the tertiary structure of the enzyme
    - Makes the reaction faster by aligning the dipoles to the neutral molecule therefore destabilizing it, in solvent they would be randomly placed so they do not interact with each other

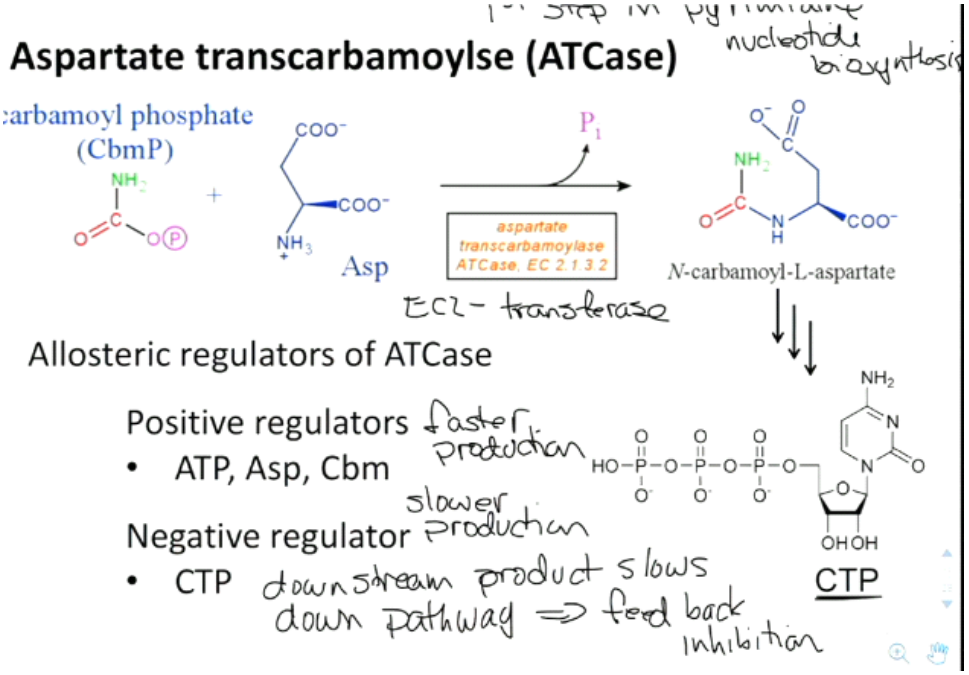


- Transition state stabilization
    - Substrate undergoing a bond cleavage event
        - "bending" - the TS - high E state
    - An enzyme that is complementary to the substrate will prevent catalysis because it stabilizes the substrate
        - Lower in energy, by binding extremely well - reaction rate plummets

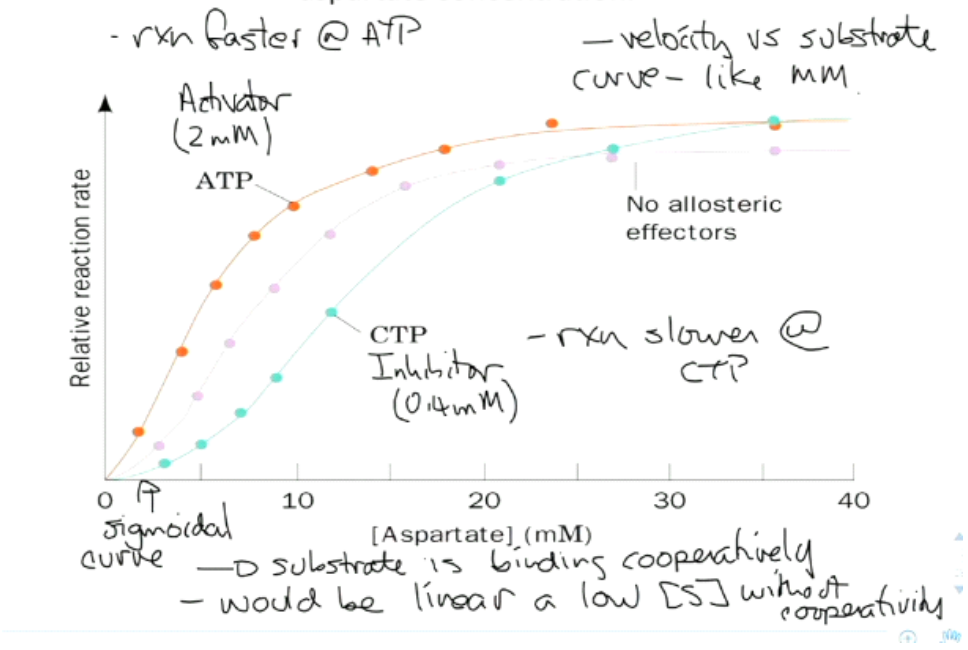
        - Lock and Key is not an accurate method - perfect binding would prevent catalysis
    - INSTEAD - it stabilizes the transition state
        - Not a great fit for the substrate, it is a bit strained in this conformation
        - Non-covalent interactions that stabilized the TS
        - The product do not look like the TS so the products come off
            - ☐ Product inhibition - if the product resemble the TS they are stabilized so much they would not come off



- Enzyme inhibitors are usually designed to mimic the TS therefore prevent the E from reacting - stablized E and inhib. complex
- Regulation of Enzymatic activity
    - Catalytic activity need to be regulated
        - Control of metabolic process, respond to the environment
    - Control the amount of enzyme available
        - Make more or less, degrade more or less
        - Long term control
        - Transcription, translation and proteolytic control
    - Control of enzyme activity
        - Allosteric control of enzyme activity
            - ☐ Rapid control - can change enzyme activity by regulation concentration of allosteric effectors - MWC or KNF model
- Aspartate ranscardamoylase (ATCase)
    - Postive regulators - ATP, Asp, Cbm
    - Negative regulator
        - CTP - the pathway ultimately makes the CTP
        - This is the rate limiting step of CTP
            - ☐ As more CTP is made, the pathway slows down
            - ☐ CTP made based on available amount
            - ☐ Feedback inhibition

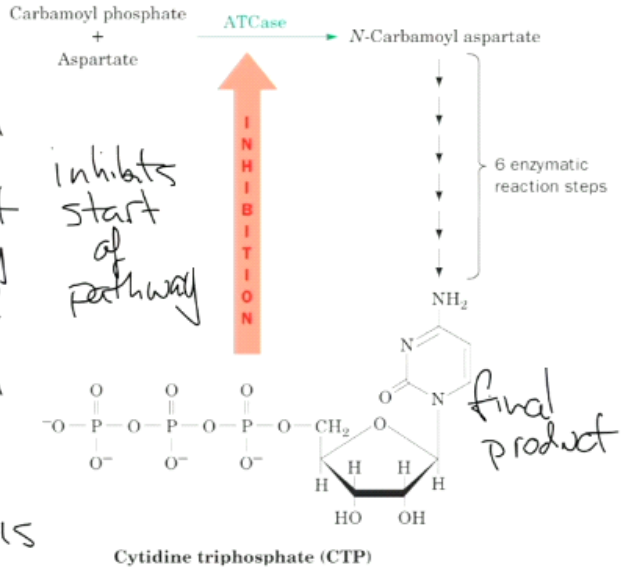


- Reaction velocity rates versus the concentration of the substrate
  - Sigmoidal phase rather than a linear phase in the CTP
  - Cooperativity curve - substrate is binding cooperatively
  - The control experiment is already a cooperative enzyme, the activator curve shows a little cooperativity



- Feedback inhibition

Feedback inhibition



- Too much product inhibits its production
- too little product and the pathway is not inhibited

inhibits start of pathway

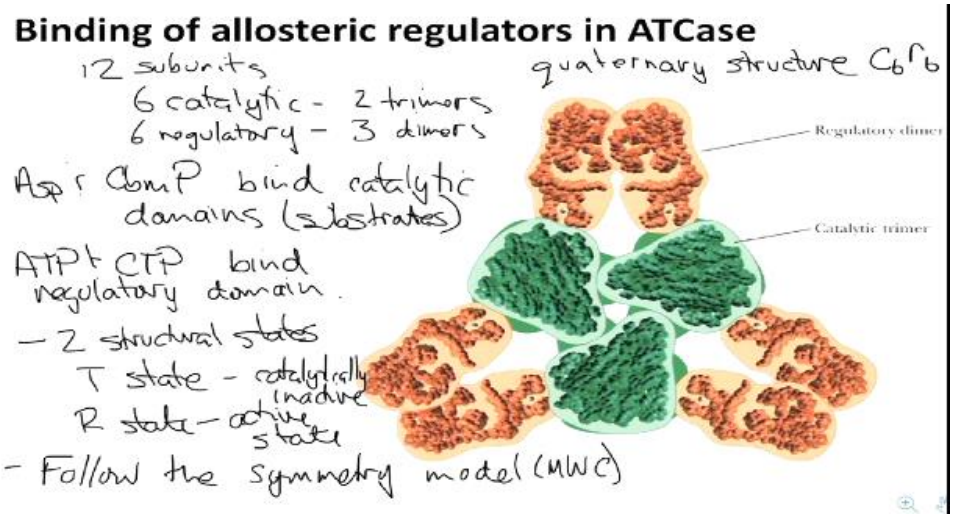
feedback inhibition helps regulate [product] thus homeostasis

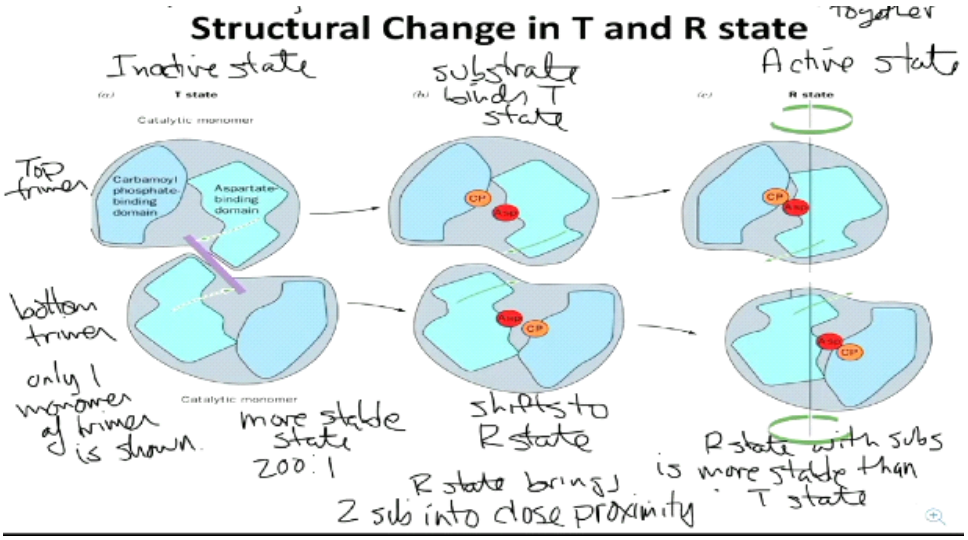
final product

Lecture 15

March-20-17 7:12 PM

- The enzymes can exist in 2 structural states
  - The T state - catalytically inactive
  - The R state - the catalytically active
    - Follow the symmetry model (MWC) - T and R state in equilibrium
    - T state is initially from stable and more abundant
      - ☐ Binding of the effector molecule would affect the equilibrium
      - ☐ Activator would cause the equilibrium to favour the R state
      - ☐ Allosteric effector molecule that is an inhibitor that would favour the T state - make it harder to reach the R state

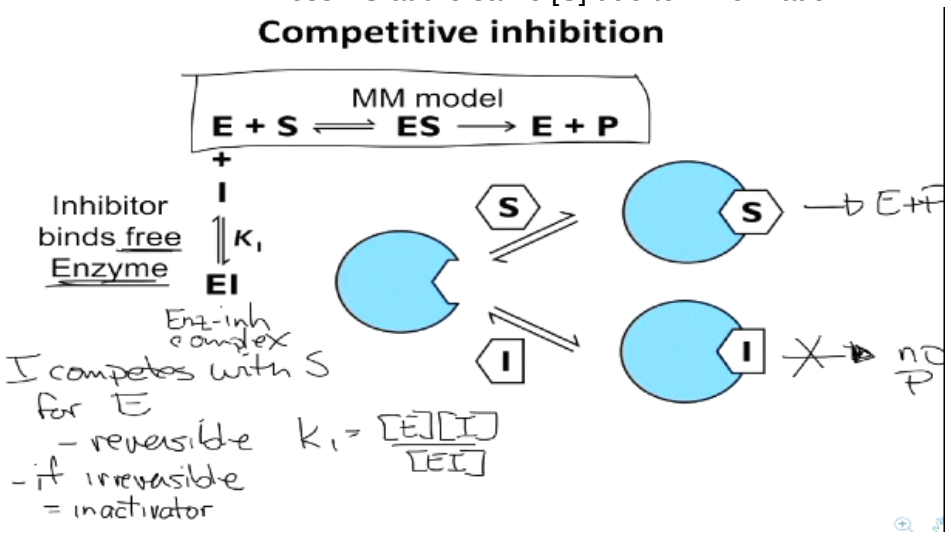


- Does not fit the sequential model, all of the subunit change conformation together
- The inactive state is more stable (200:1) - in the T state the Carbomoyl (cbmp) and Asp are very far
  - R state brings the binding sites closer together
  - Substrate binds T state
    - Enzyme undergoes a conformational shift which forms the R state where the CBMP and Asp are close together
    - Active state
- 

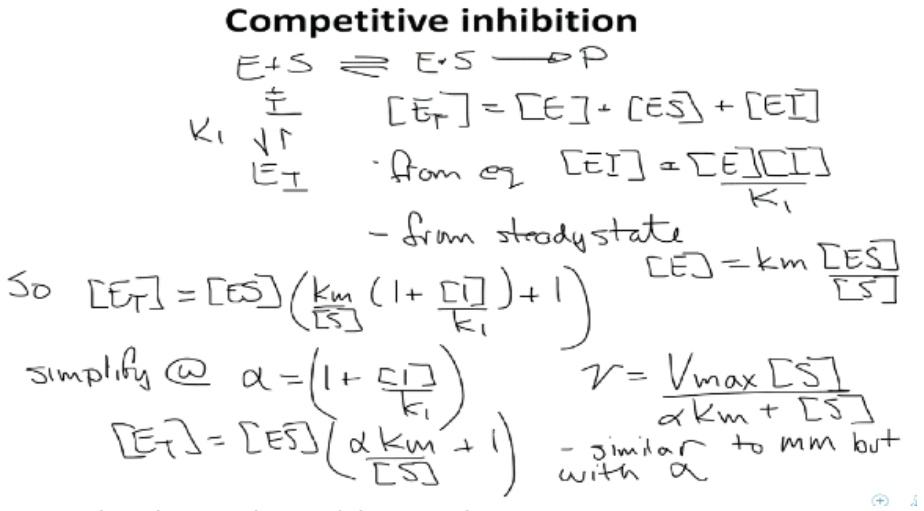
- Non-substrate can bind to the regulatory or catalytic domain to cause a shift between T and R state

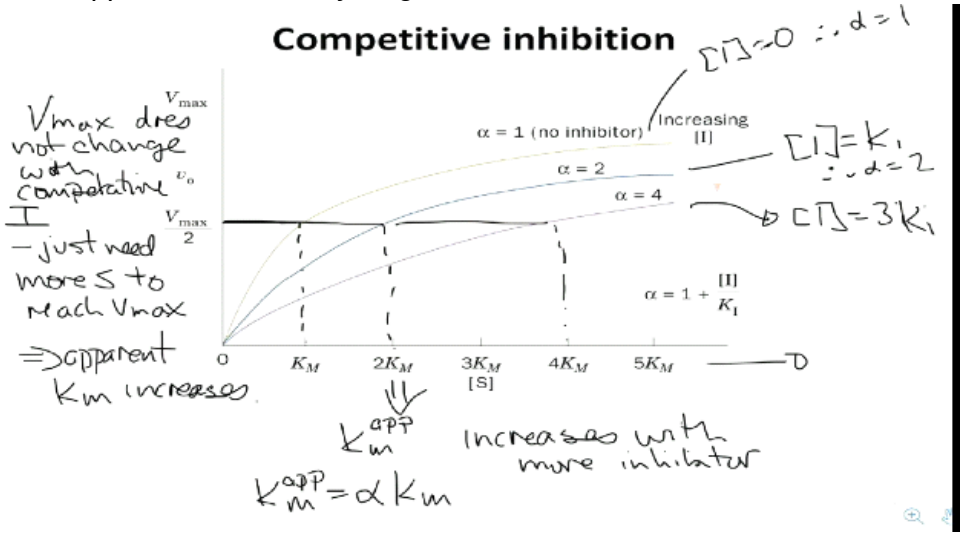
- Regulatory domain effects T and R state
  - CTP binding stabilizes the T state and prevents the enzymes from entering the

equilibrium with the R state
    - ATP can bind to R state to stabilize it, pulling it out of the equilibrium, due to Le Chatelier principle forms more R state
        - Since it is already in the R state there is no cooperativity because already high affinity is achieved
        - CTP binding to the T state shows greater cooperativity
            - ☐ Substrate binding forces the high affinity conformation
            - ☐ Much larger difference in the equilibrium to reach high affinity - from the T state to the R state
- Allosteric theory of Enzyme regulation
    - Activator will bind preferentially to the R state
        - Active conformation or high cooperativity
    - Inhibitor will bind preferentially to the T state
        - Inactive conformation or low affinity state
    - Phosphofructokinase is an example
- Inhibition of enzymes
    - Activity of an enzyme can be slowed by binding of effector
    - Seen with CTP binding to ATCase
    - 3 mechanisms for inhibition
        - Competitive
        - Uncompetitive
        - Mixed
- Competitive Inhibition
    - Inhibitor binds free enzyme
    - Inhibitor binds to the free enzyme to make an EI complex (Enz - Inh)
    - Inhibitor is competing with substrate - E binds either the substrate or inhibitor
        - Reversible arrows of the inhibitor binding to the enzyme
        - Reversible inhibition - the inhibitor can reversibly bind with the inhibitor
        - $K_I = ([E][I])/[EI]$ - written as a dissociation constant
        - Some inhibitor bind irreversibly - inactivator
        - Less free enzyme so less enzyme substrate complex - therefore slows down the rate of the reaction
            - ☐ Product formation is k_{cat}=[ES]
            - ☐ Less ES at the same [S] due to EI formation

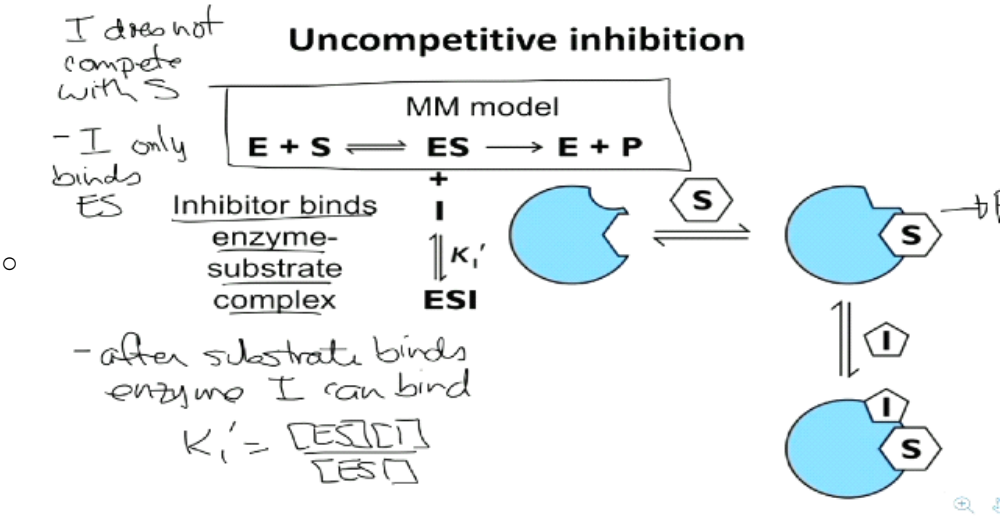


- The enzyme substrate complex is really important to determine the velocity of the reaction
- To achieve final equation for velocity,
    - Isolate for [ES] then multiply by k_{cat} on both sides, and solve for $k_{cat}*[E_T]$ = Vmax and $k_{cat} * [ES] = v_{cat}$

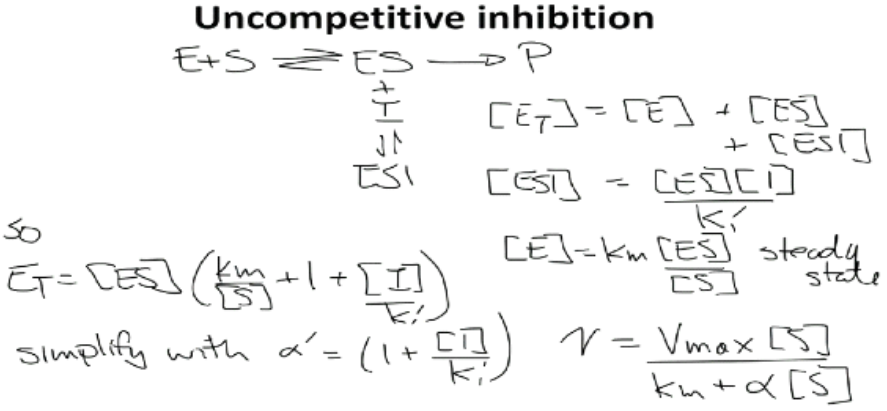
- 

- It impacts the observed K_M of the reaction
- This curve should continue up to V_{max} because as the substrate concentration increase it can prevent binding to the inhibitor, with enough we can achieve enzyme saturation and have just free inhibitor
    - Therefore the apparent K_M should be at 1/2 V_{max}
    - Apparent K_m is always higher than the true K_M
- 

- Uncompetitive inhibition
    - Inhibitor binds to the enzyme-substrate complex
    - I does not compete with S
        - I only binds to the enzyme substrate complex
        - Reversible inhibition
            - ☐ Ternary complex (ESI)
            - ☐ $K_I' = ([ES][I])/[ESI]$
    - Enzyme has two different binding sites in order to allow for binding at the same time



- Uncompetitive inhibition derivation

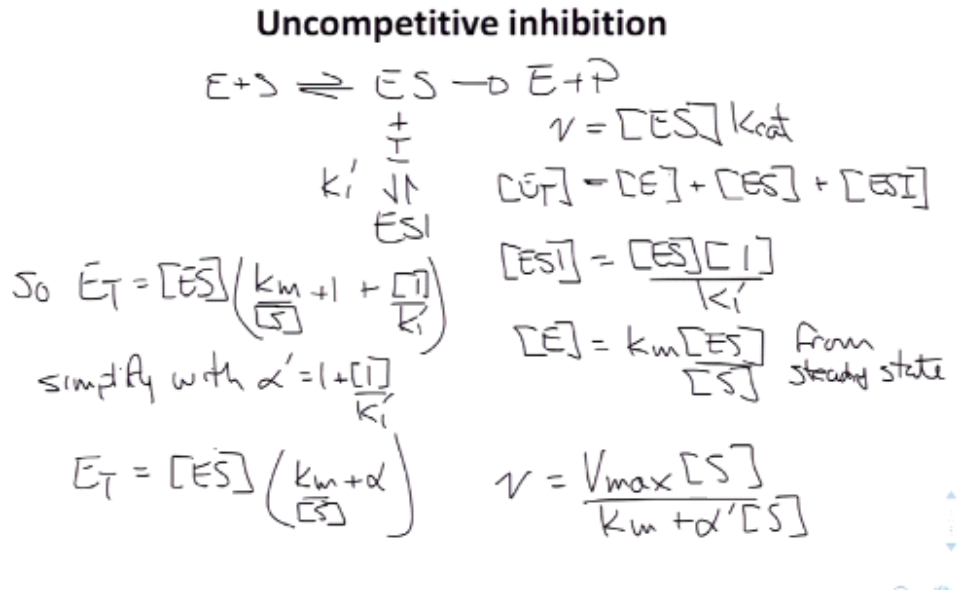


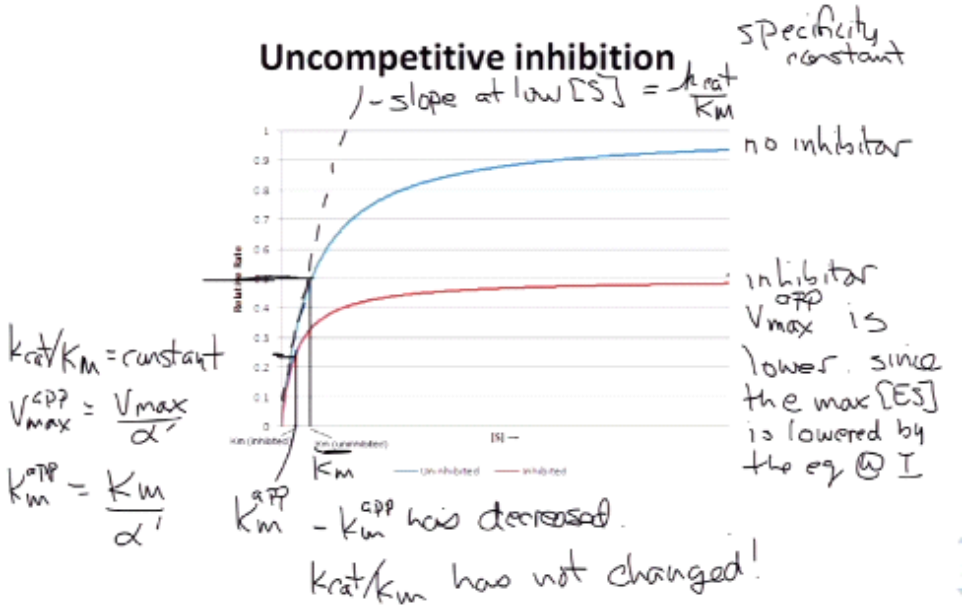
-

Lecture 16 - Inhibition

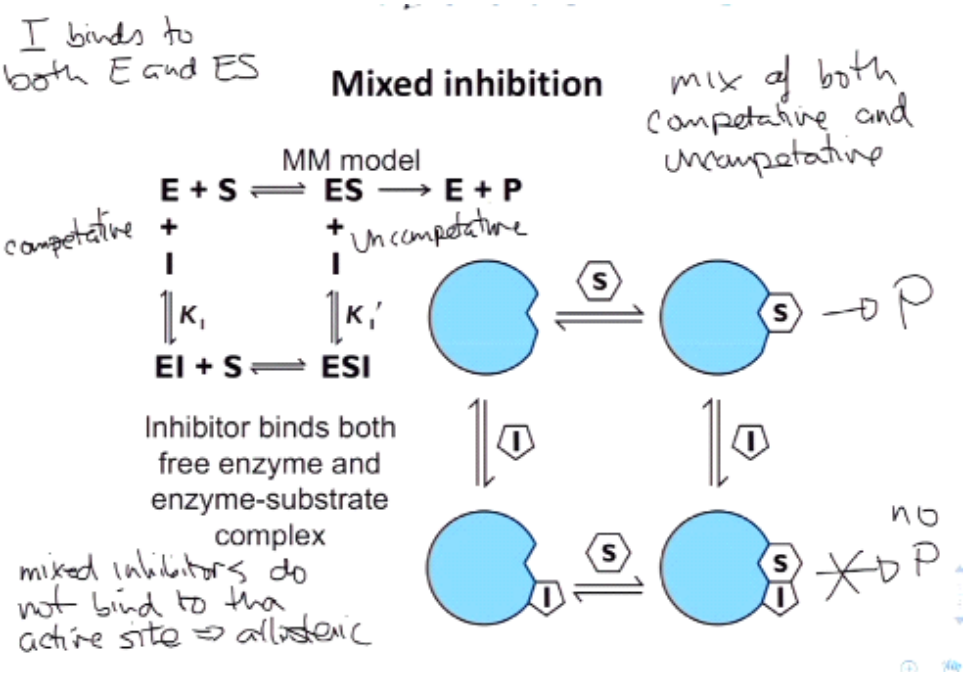
March-20-17 9:07 PM

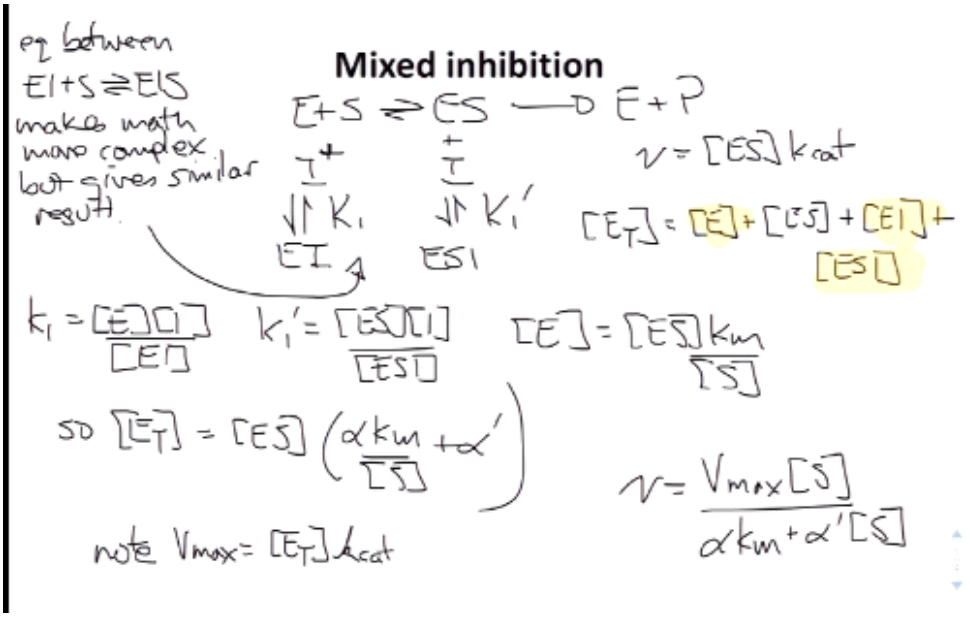
- Uncompetitive inhibition
  - V = [ES]k_{cat}



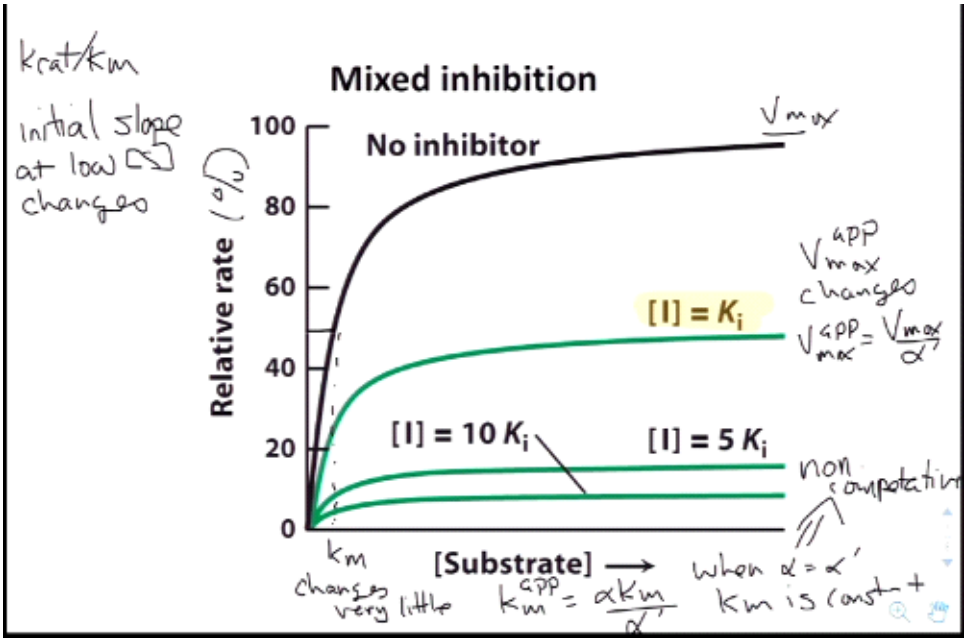
- The linear relationship when below Km is essentially just Km so a direct linear relationship
  - At low [S] = k_{cat}/k_M
    - Specificity constant
  - Tells you the energy between the substrate and the products and the competitive between substrates for the enzyme
  - Decreases the V_{max} called the V_{max} apparent, it is lower than the actual V_{max}
  - The enzyme is never fully able to be saturated as a certain number of enzymes are always unable to create products due to the inhibitor so V_{max} decreases
    - K_M will also be smaller (since K_m= 1/2 V_{max})
  - K_{cat}/K_M has not changed - this is constant at low S conc.
  - 

- Mixed
  - The inhibitor binds other than the active site





- Decrease in V_{max} - enzyme substrate complex can be removed, so full V_{max} cannot be reached, V_{max} apparent is lower than actual value
- K_m changes very little, $K_{m\ app,}$ when alpha and alpha ' are equal equal likelihood of binding of inhibitor to the E or ES complex



Difference between inhibition mechanisms

Type of Inhibitson	V_{max}^{app}	K_M^{app}
None	V_{max}	K_M
Competitive	V_{max}	αK_M
Uncompetitive	V_{max}/α'	K_M/α'
Mixed	V_{max}/α'	$\alpha K_M/\alpha'$

$\alpha = 1 + \frac{[I]}{K_I}$ and $\alpha' = 1 + \frac{[I]}{K_I'}$

- K_M increases in the competitive model
- In the mixed model - K_m is likely to stay constant

Lecture 17

April-23-17 10:59 AM

Nucleic acids: diverse roles in biochemistry

- polymers of nucleic acids: DNA RNA

  Role - information storage - DNA
  - gene expression - RNA
  - can be catalytic - RNA in ribosome

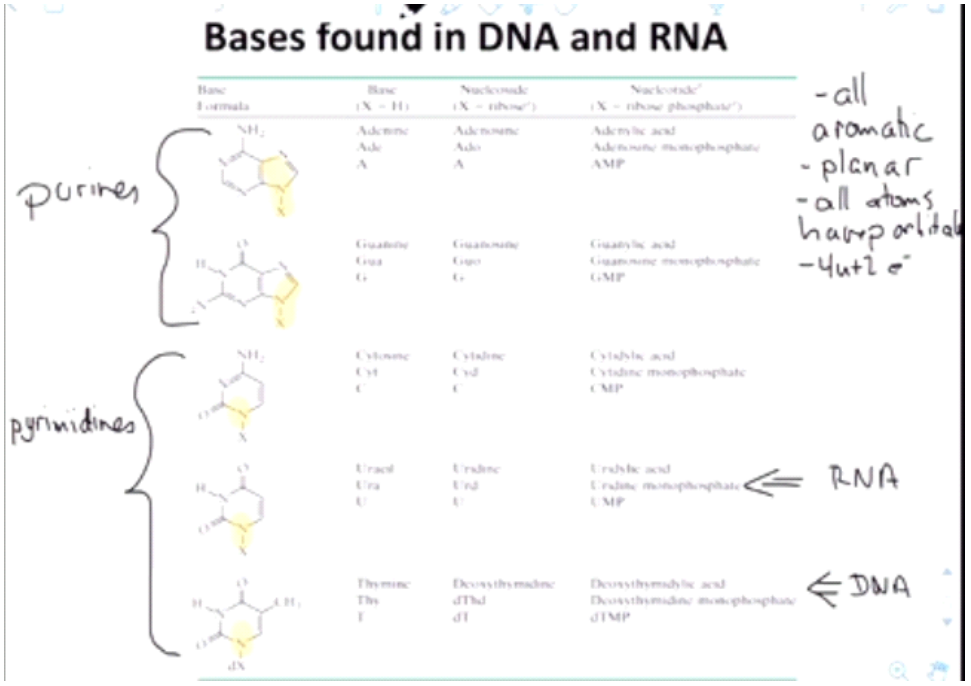
- monomers: nucleotide triphosphates

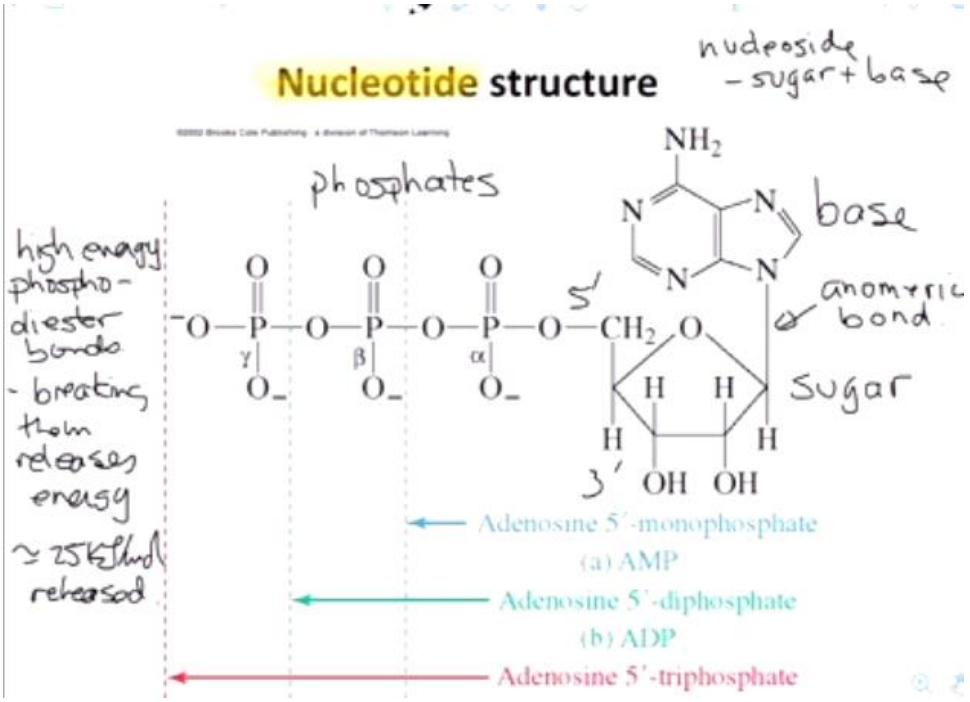
  ATP GTP, CTP etc
  - provide energy for metabolic processes
  - allosteric regulators eg ATCase

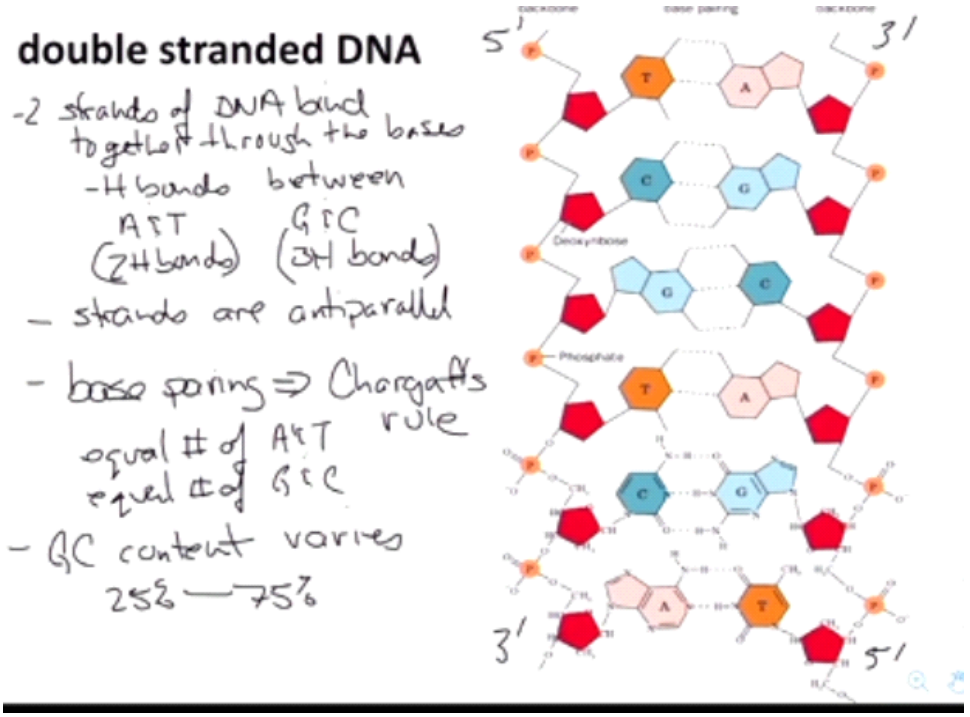
- cofactors: NADH, FAD, CoA,

  look up these structures
  - identify the nucleic acid parts

- **Bases found in DNA and RNA**



- 

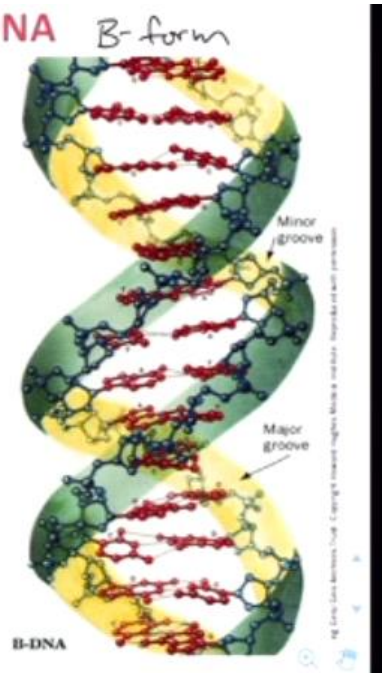


- Planar interaction between bases ensure that th hydogen bonding is as strong as possible

Double helical DNA B-form

- Structure determined by Watson & Crick in 1953

- right-handed helix
- **20** Å diameter double helix.
- Bases occupy the core and sugar-phosphate chains are on the outside.
- Planes of bases are perpendicular to the helix axis.
-
- Each bases is hydrogen bonded to opposite strand to form a planar base pair.



- Pitch (rise per turn) = 34A (n=10)
- 36 degree per bp
- 3.4 A distance between bases
- A - DNA - 23A - dehydrated form
  - 11bp/turn
  - Thickness - 2.4A/base
- Z-DNA - left handed helix - very rare
- Watson-Crick base pairing

Watson-Crick base pairs

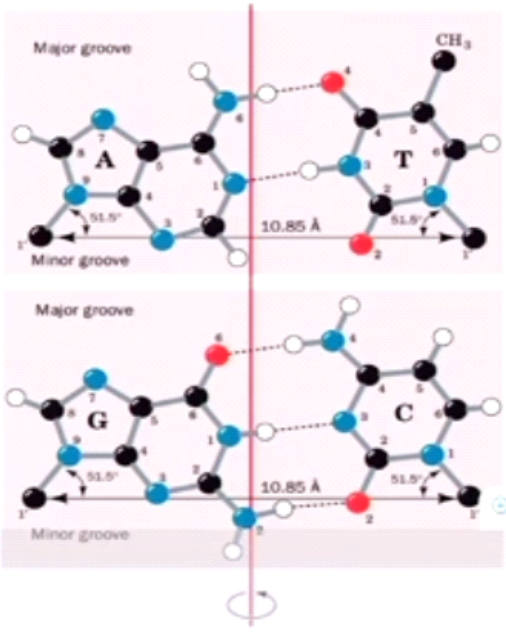
Adenine pairs with thymine.
Guanine pairs with cytosine.

base pairs are interchangeable, exchange does not alter the positions of the sugar phosphate backbone.

The top edge of each base pair is structurally distinct from the bottom edge.

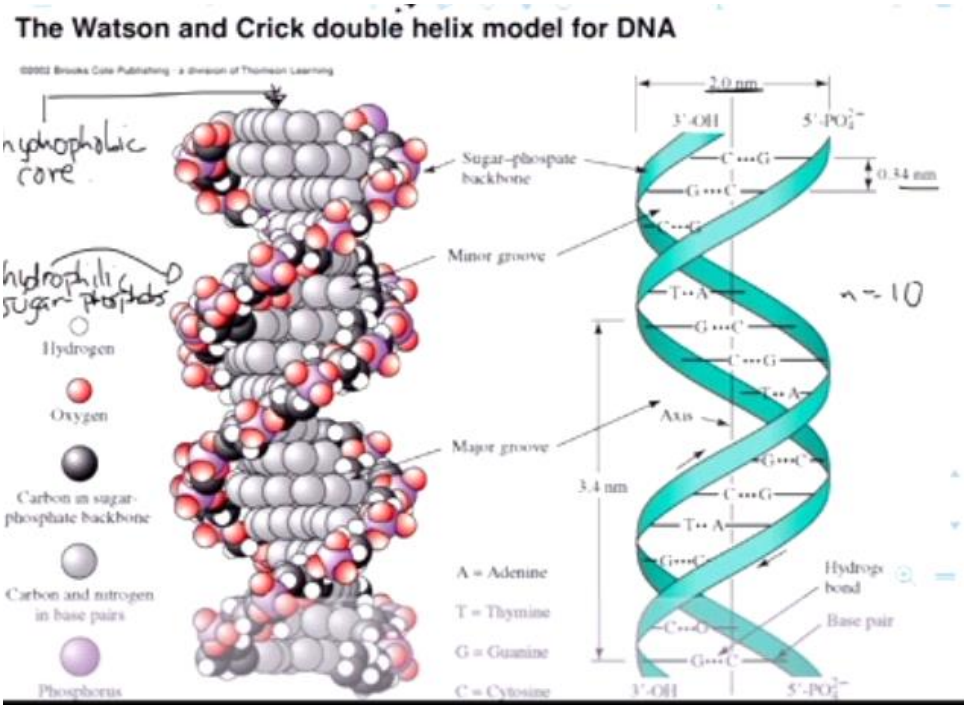
Minor groove exposes the edge from the C1' atom (open toward bottom)

Major groove exposes the opposite edge of each bp.



- Distance between the GC and AT are the SAME (10.85A)
- The sugar backbone STAYS the SAME, regardless of what bases are pairing in the middle
- The structure of B-DNA is SEQUENCE INDEPENDENT
  - The sequence does not determine the structure
- The spacing between the phosphate backbone
- Major groove in the top part of the A and G, minor groove is the bottom part of it
- The core of DNA is hydrophobic - van der waal packing
  - Excluding water in the center - very tight packing

- Hydrophillic sugar phosphate groups
  - Solvates the DNA

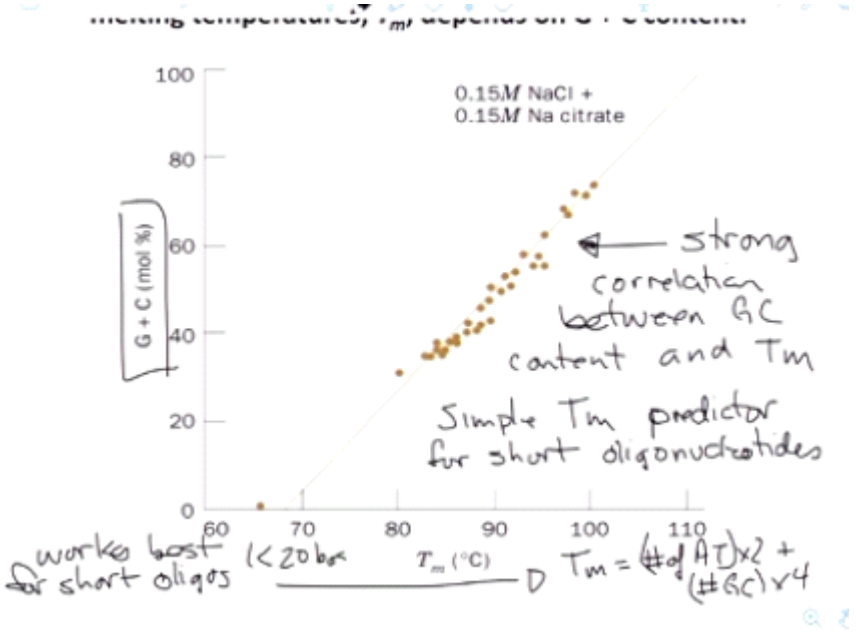


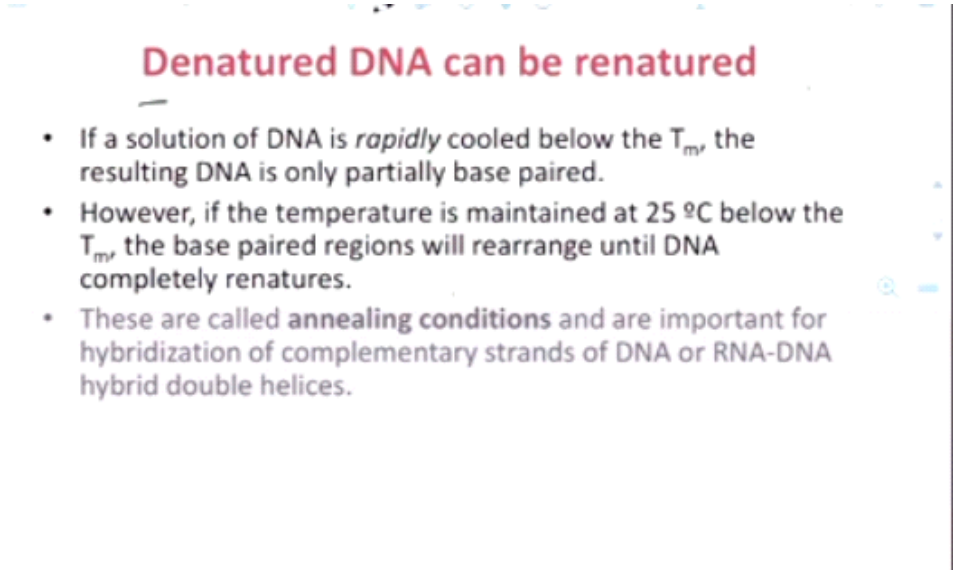
2 key forces:

1. **H-bonds between complementary bases on opposite strands:**
   - **2 H-bonds in A-T pair**
   - **3 H-bonds in G-C pair** – stronger force holding GC together than AT
2. **Van der Waals forces and hydrophobic interactions between "stacked bases"**
   - **Aromatic bases have π-electrons that interact via attractive Van der Waals forces.**

π-stacking π e⁻ are polarizable – lots of van der Waal forces

- Denaturation to renaturation
  - Forms a random coil - because the structure is in an atypical structure
  - Cooling it down causes renaturation - very reversible to reform the original DNA strand with correct base pairing
- Denaturation is cooperative
  - Initial destabilization allows entry of water which further destabilizes the structures, causing the DNA to unzip
  - The midpoint of the graph of double stranded to single stranded is the melting temperature
    - Depends on solvent, water or methanol
    - Concentration and type of ion - can have counter ions
    - pH - protonation of the phosphate backbone
    - GC content

- 



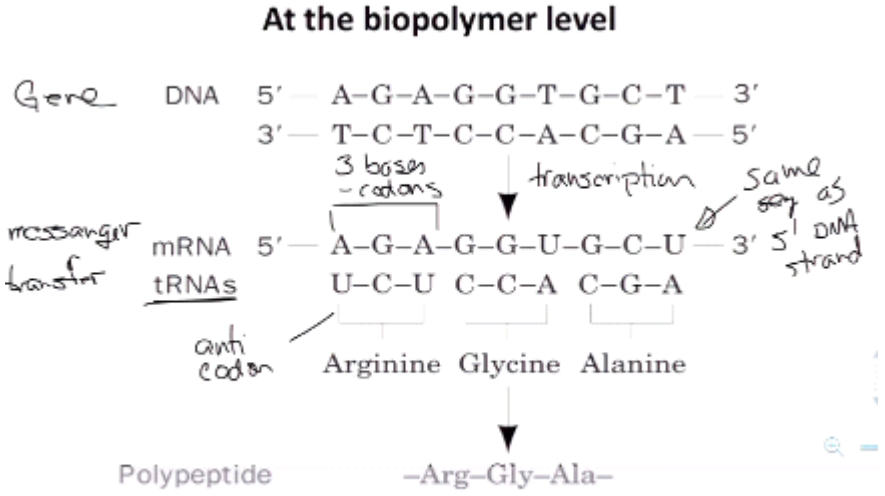
25 degress Celsius allows reaarrangment to occur to reach lowest energy structure.

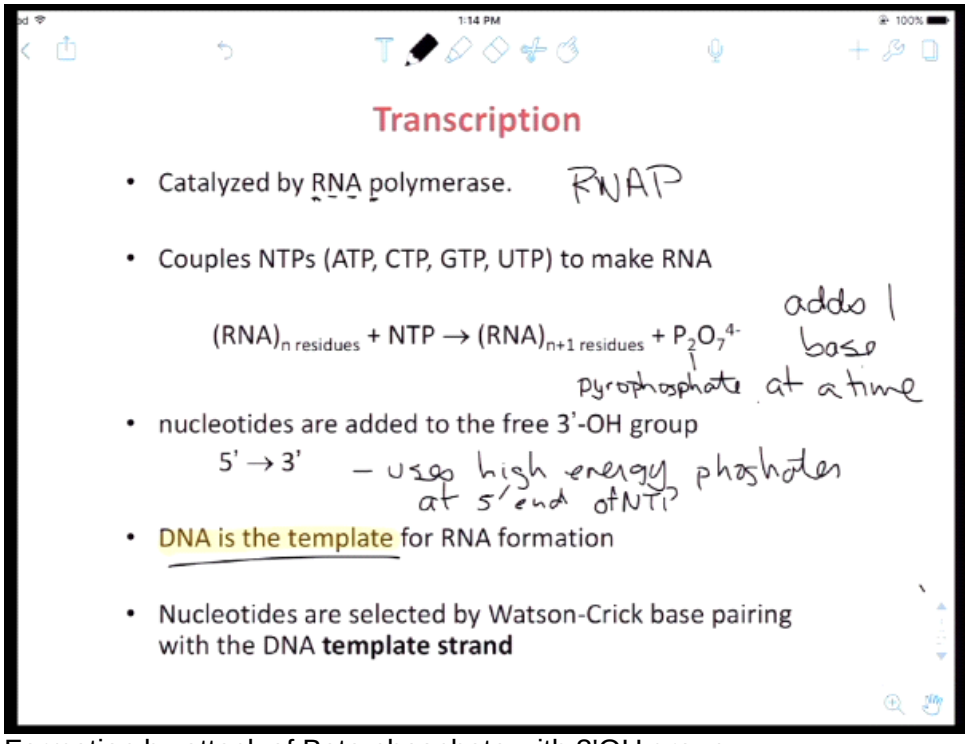
- Important for hybridization
- Rapid cooling traps the DNA in non-helical structure, partial renaturation
- Partial renatured DNA
    - Trap states leading to local good base pairing but not fully stabilized
    - If cooled too much, no energy available to drive rearrangement to reach a thermodynamic minimum
    - Rapid cooling prevent equilibration
    - Intermolecular aggregation and intramolecular aggregation
        - Inter - double stranded interaction
        - Intra - single strand interacting on its own
- Contour length (um) - DNA stretched from end to end
    - Greater than 20,000 base pairs, incredibly sensitive to mechanical stress
        - Chemical bonds are BROKEN - cannot reform again

Lecture 18

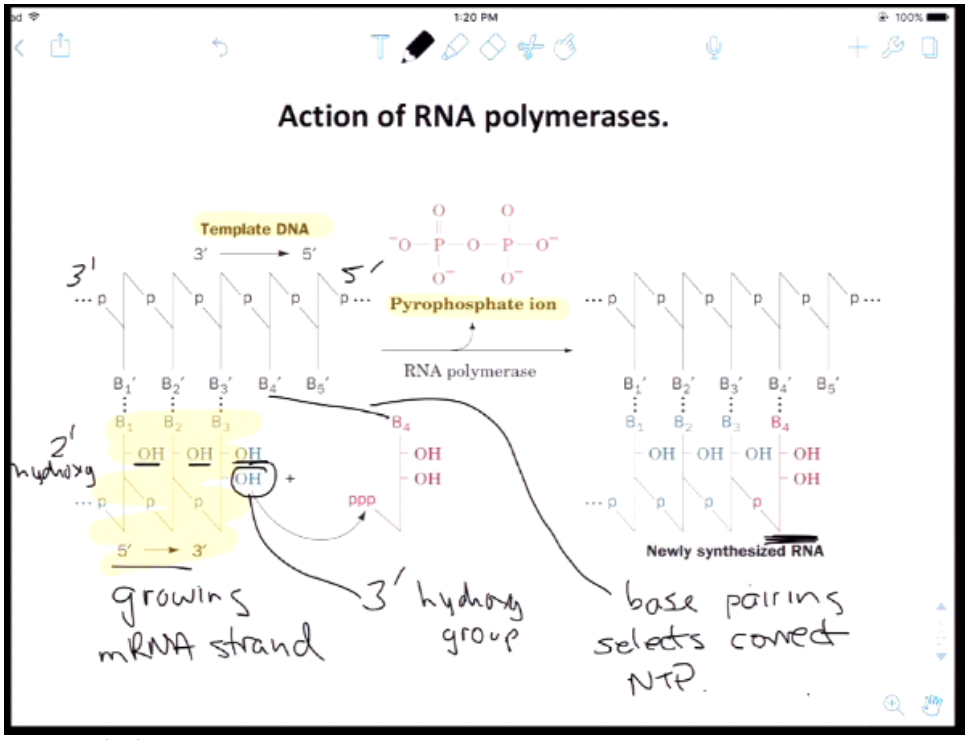
April-23-17 12:38 PM

- 5' strand of mRNA is developped by complimenting the the 3' strand of DNA, producing an exact copy of the 5'strand of DNA except with U instead of T
- TRANSCRIPTION - READ 3'to 5' TRANSLATION - READ 5' to 3'
- 3 base segments - codons
  - Complementation with tRNA's - transfer
    - The anti-codon complements the codon on the mRNA
  - FROM N TERMINS TO THE C TERMINUS





- Formation by attack of Beta phosphate with 3'OH group

- 

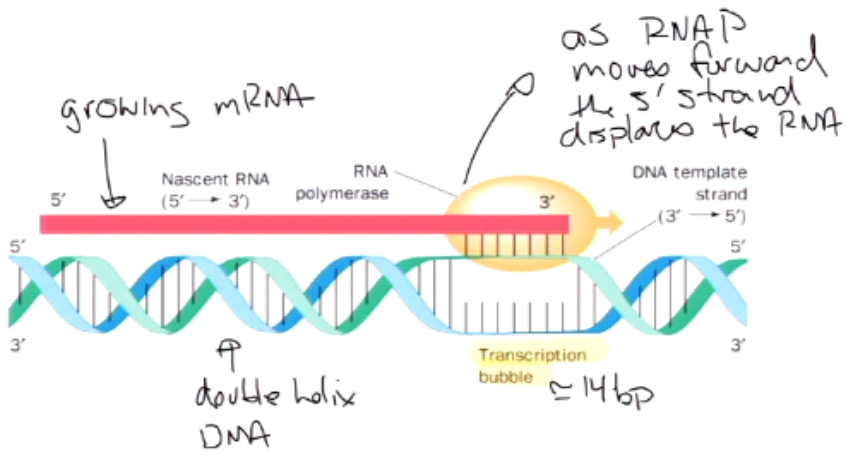
- Transcription

Transcription

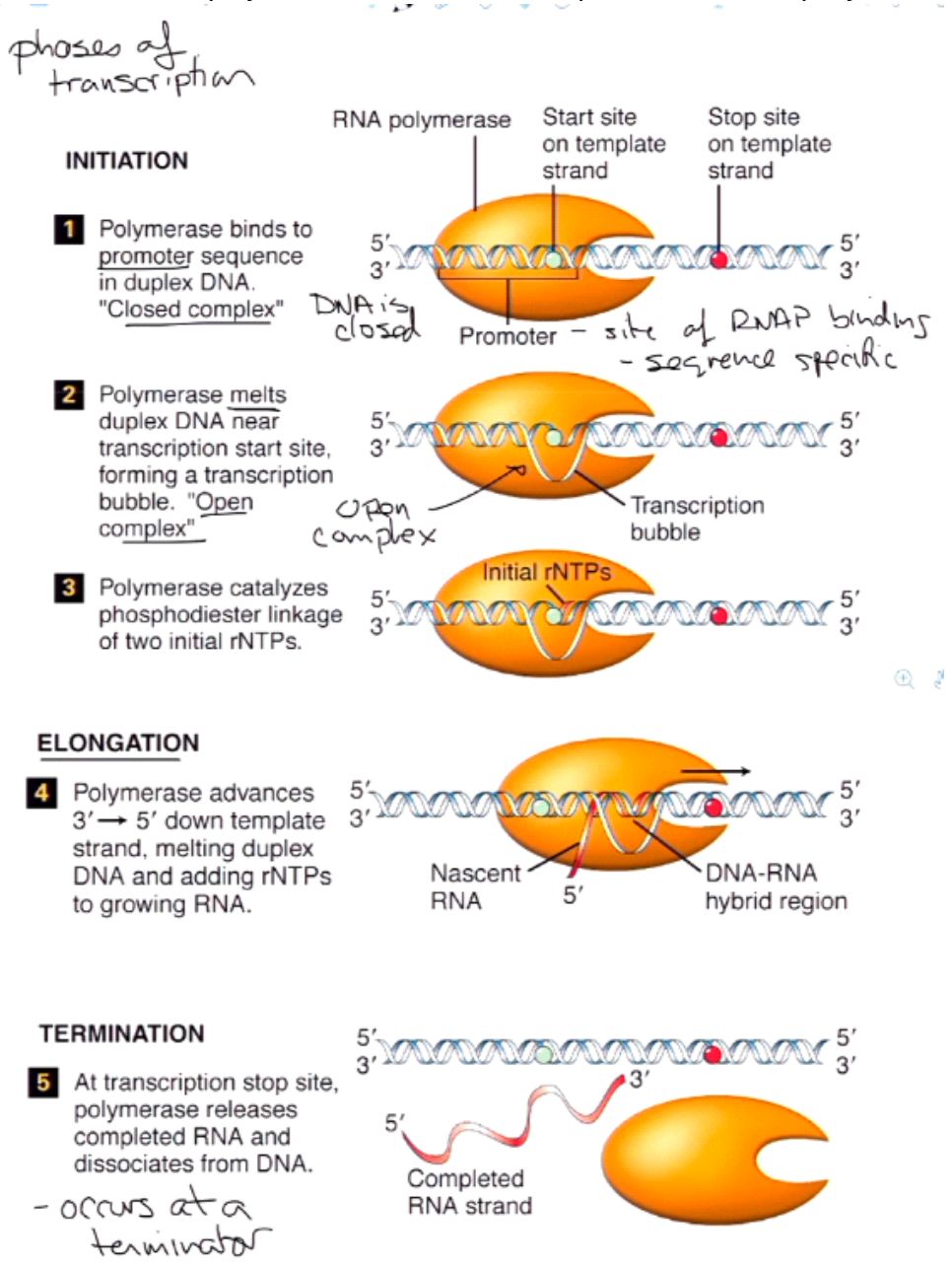
- As RNA polymerase moves along the duplex DNA it creates a **transcription bubble**
  - ≃ 14 bp that are not paired
  - site of RNA synthesis
- This forms a short DNA-RNA hybrid with newly synthesized RNA.
  - DNA-RNA hybrid is a double helix with antiparallel strands
- DNA template strand is read 3' → 5'
  - this makes a RNA copy of the 5′ DNA strand.

- As RNAP moves away, the DNA comes in and displaces the mRNA and closes up the DNA as RNAP moves along - 5' strand in NOT duplexed by RNA, it is the 3' strand

Function of the transcription bubble.



- Phases of transcription
  - Inititation - closed complex - bubble not formed because mRNA is needed to open the DNA - DNA is closed
  - Promoter sequence is sequence specific
  - Transcriptional start site- site of 3 rNTP binding
  - DNA polymerase cannot initiate replication like this (they need an RNA primer)



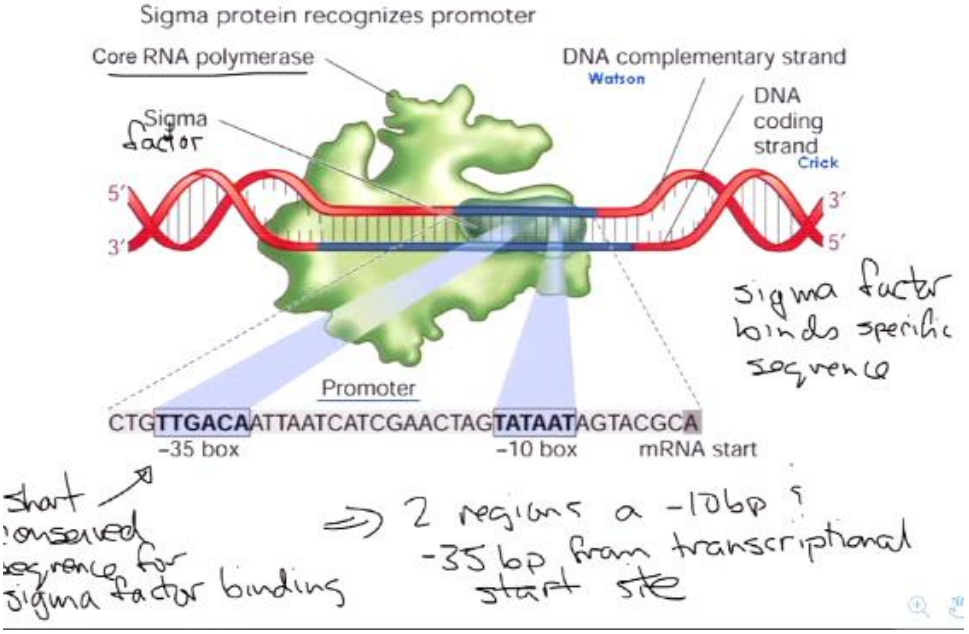
- RNAP complex is released from dsDNA at terminator

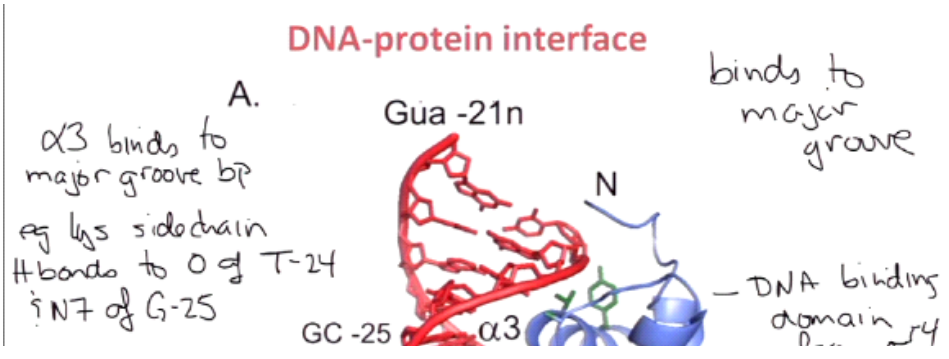
Transcription

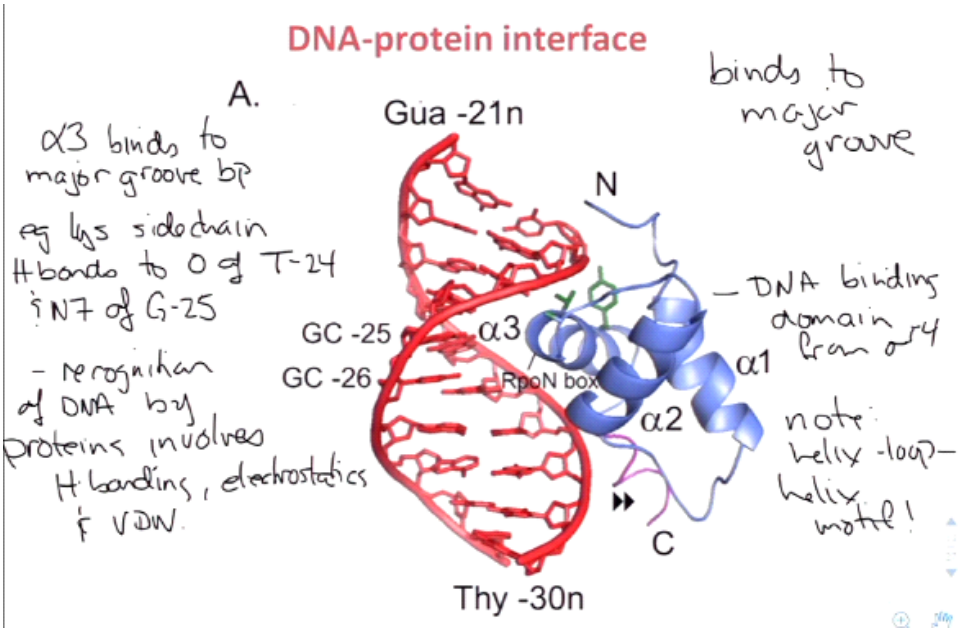
- DNA contains control sites that specify where the RNA polymerse initiates transcription.
  - sequence specific sites
  - promoters
- **activators** and **repressors** control the sites in prokaryotes.
  - activators & repressors bind DNA sequence specifically
  - contact major or minor grooves.
- **Transcription factors** bind to these sites in eukaryotes.
  - transcriptional regulation
  - main control mechanism for protein levels in cells.
- **Promoter**-DNA sequence that precedes the transcriptional initiation site. 5' to the coding strand.

- Bacterial RNAP assocaited with a sigma factor - specific protein that binds the promoter sequence of DNA
- CORE RNAP binds to the sigma factor bound to the promoter
  - Interact at 2 different sites - sigma factor is usually a dimer

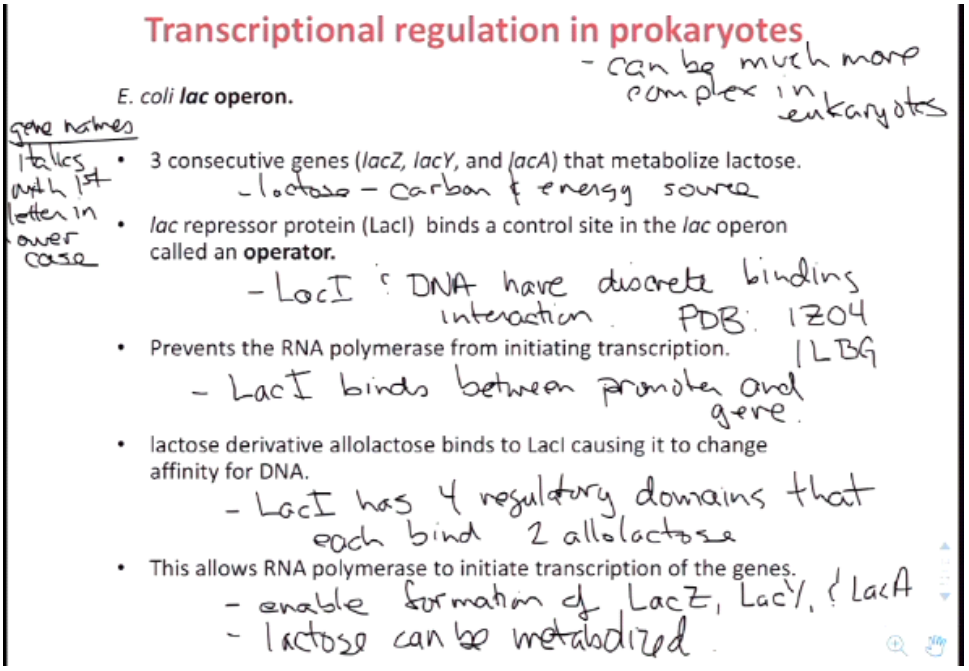
Bacterial RNAP is directed to DNA by sigma factor

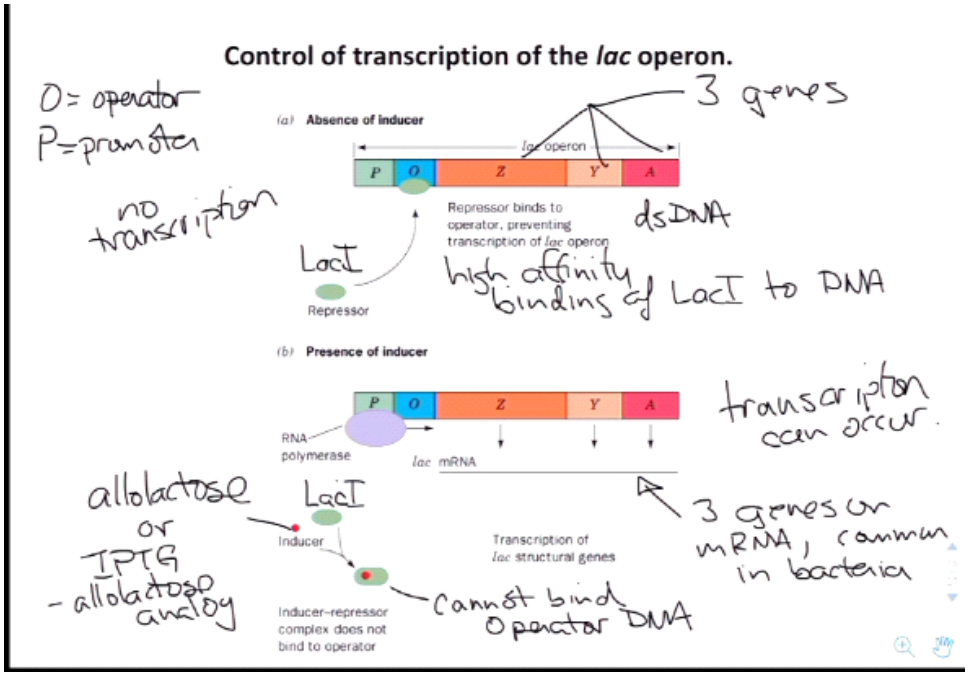


- DNA binding domain at the major groove
  - Helix loop helix motif (alpha - alpha motif)
  - Alpha 3 interact with the bases in the major groove
  - Lys sidechain H bonds to the O of the T-24
    - N7 of G - 25
  - 

- 

- Genes encodes a protein
  - Genes are ALWAYS written in italics - DNA
  - Protein names are not italicized



- 
Control of transcription of the lac operon.
O= operator
P=promoter
3 genes
(a) Absence of inducer
lac operon
P
O
Z
Y
A
no transcription
Repressor binds to operator, preventing transcription of lac operon
dsDNA
LacI
Repressor
high affinity binding of LacI to DNA
(b) Presence of inducer
transcription can occur.
RNA polymerase
lac mRNA
allolactose or IPTG - allolactose analog
LacI
Inducer
3 genes on mRNA, common in bacteria
Transcription of lac structural genes
cannot bind Operator DNA
Inducer-repressor complex does not bind to operator

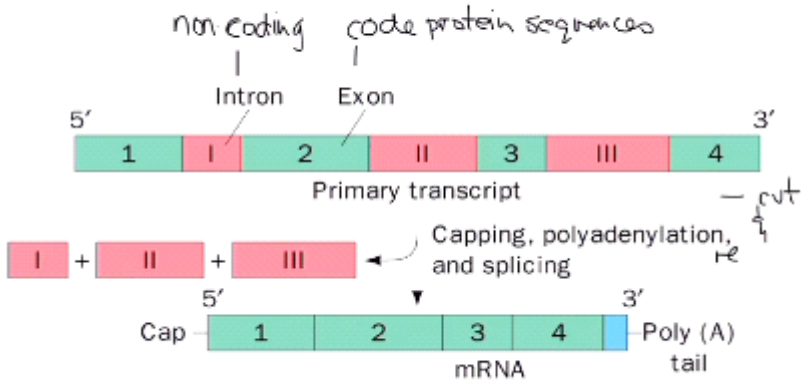
Lecture 19

April-23-17 2:44 PM

- mRNA is cleaved at sites where introns are and stitched back together

Gene splicing!

Post-transcriptional processing of eukaryotic mRNAs.



- In prokaryotes, transcription and translation both take place in the cytosol. (No nucleus)
- Prokaryotic mRNAs have a short lifetime (avg. 1-3 min). They are degraded by **nucleases**.
- Rapid turnover in prokaryotes allows the prokaryote to respond quickly to the environment.
  - fast response → doubling time of E. coli 15 min
- In eukaryotic cells, RNAs are transcribed and ~~post translationally modified~~ in the nucleus, then sent to cytosol. processed.
- Eukaryotic mRNAs have lifetimes of several days.

Translation: Protein synthesis

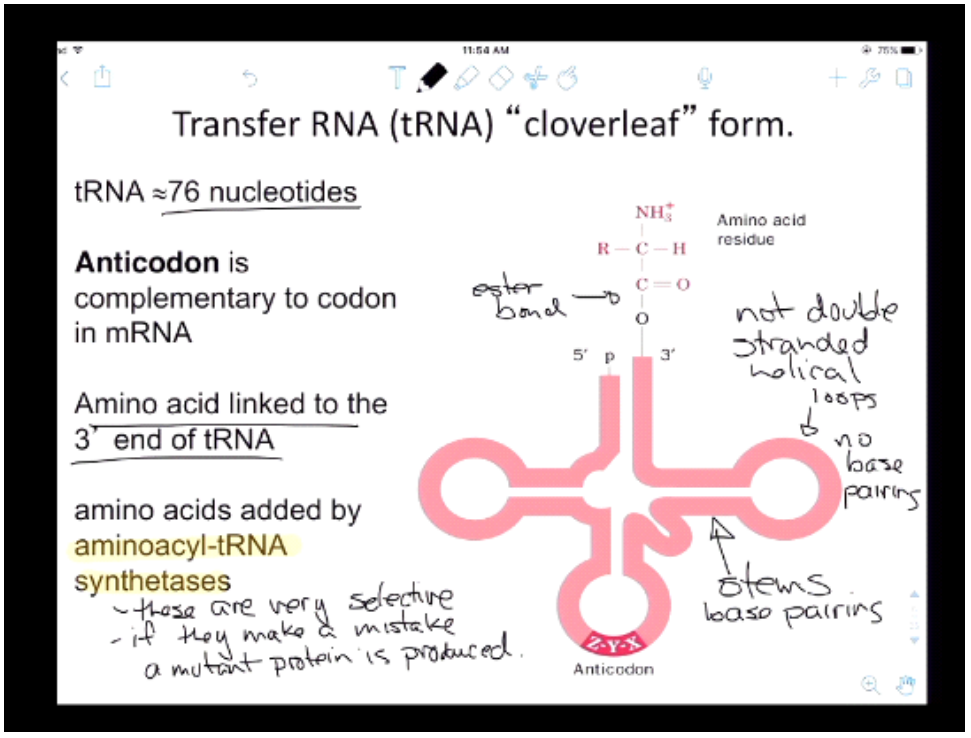
- Polypeptides are synthesized from mRNA by **ribosomes.**
  - bacterial ribosome ≈ 2,500 kDa
  - eukaryotic ribosome ≈ 4,300 kDa
  - Very big
- Ribosomes - ribosomal RNA and protein.
  - 2/3 RNA 1/3 protein
  - rRNA
- mRNA sequences contain **codons**

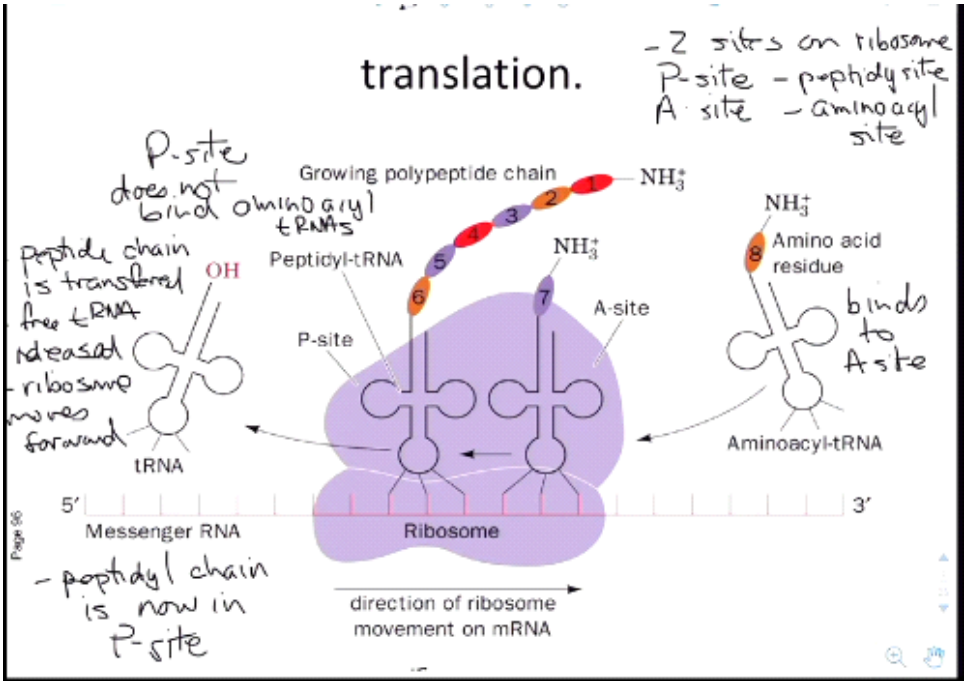
  Codon = consecutive 3-nucleotides that specify an amino acid.

- **Transfer RNAs (tRNAs)** deliver amino acids to the ribosome. have aa covalently attached

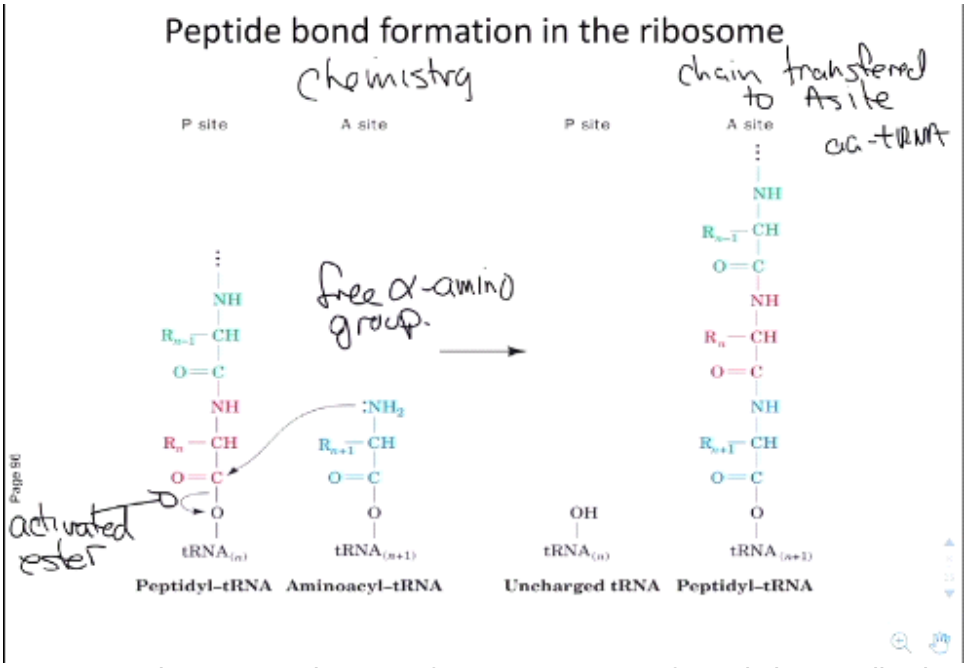
  Have anti-codons – 3 complementary bases to codon
  Watson-Crick base pairing selects correct tRNA
- mRNA binds ribosome and tRNAs bind mRNA to template protein formation

Anti-codon is found in loop to ensure proper binding, aminoacyl-tRNA covanlently attached to the 3' alcohol group

- Carboxylic acid on amino acid turned into an ester therefore activating the amino acid - fidelity of translation is the right amino acid is added to the right tRNA
    - Can result in a mutation
- Aminoacyl tRNA synthetase - incredibly sensitive to the correct anti-codon nad amino-acid
- 



- A 0site binds to tRNA - it liks the positive charge - amino group
- P-site prefer a neutral charge - amide group



- Formyl group on nitrogen of Met-tRNA, so no formal charge allowing it to go directly to the P-site

Genetic code

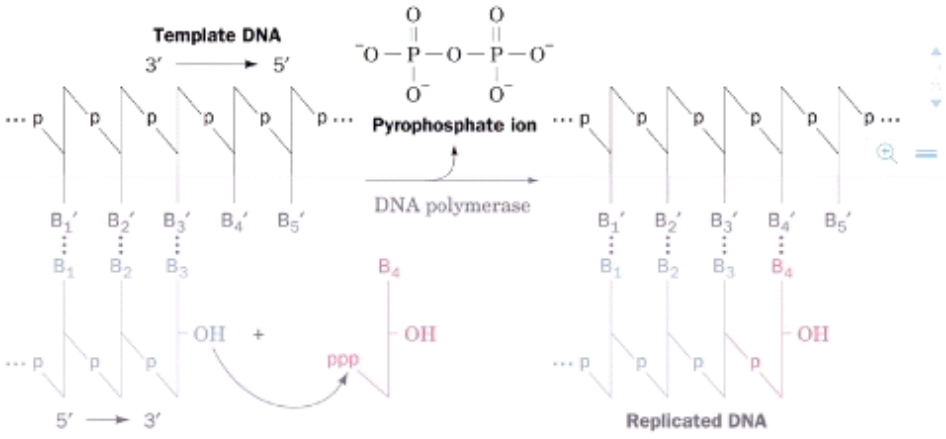
- sequence of bases in a codon specify an amino acid
  - nearly universal what aa are encoded by what codons
- 4 possible bases (U[T], C, A, and G) 3 positions
  - $4^3 = 64$ codons
- 61 codons specify amino acids, 3 (**UAA, UAG,** and **UGA)** are **stop codons**
- All AA but Met, Trp are specified by more than one codon.
  - most aa have codons that are synonomous
- Leu, Ser, Arg are specified by six codons.
  - most common aa have many codons.
- Translation is initiated at the AUG codon (Met) but this tRNA differs from the tRNA for internal amino acid the Met codon.
  - all genes start with AUG ⇒ Met ATG in DNA

- Wobble pairing - synonymous mutation is the wobble pair - can still code for the same amino acid
- 3 reading frames in mRNA and 6 reading frames in DNA

- DNA is replicated similar to RNA with some differences
  - uses dNTPs - deoxyribose
  - call DNA polymerase
- RNA polymerase can link together two nucleotides but DNA polymerase can only extend ~~in the 5' to 3'~~ oligonucleotides
  - catalyzes synthesis 5'→3' like RNAP
- DNA polymerase needs an oligonucleotide **primer** to initiate synthesis
  - RNA acts as the primer
  - ~10bp made by Primase in E coli
  - RNA is used so it can be distinguished later on and error correc

DNA polymerase

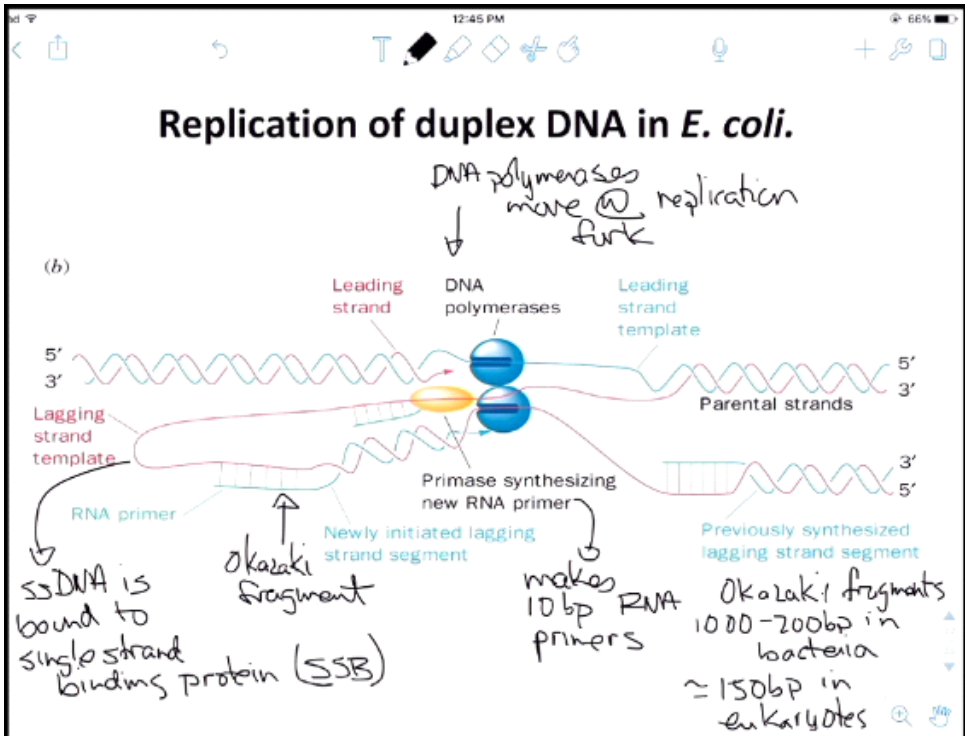
templo



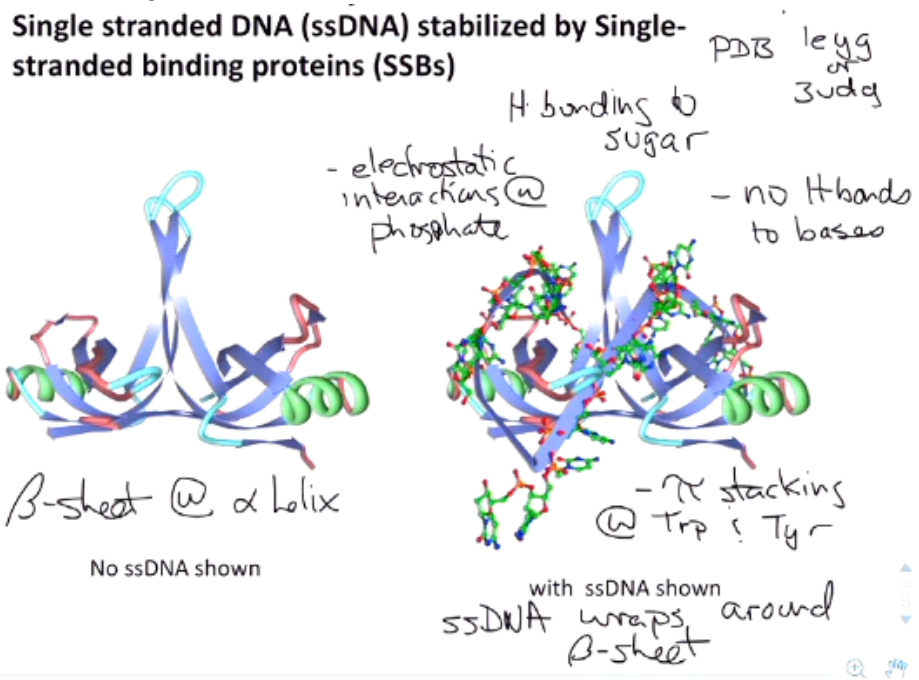
DNA strands replicated in different ways

- DNA strands are simultaneously replicated.
  - 2 strands in DNA one 3'→5' one 5'→3'
  - each is a template for one DNA polymerase
- **replication fork** – 2 parental DNA are pried apart and the 2 daughter strands are synthesized.
  - double helix is melted to access single stranded DNA (ssDNA) as template for synthesis
- **Leading strand** -- 3' to 5' parental template continuously synthesized in the 5' to 3' direction
  - replicated easily, just like RNAP activity
- **Lagging strand** – 5' to 3' strands is discontinuously synthesized
  - synthesis on lagging strand is more complex

- Lagging strand - DNA binds to single strand binding protein (SSB)



- Sequence independent replication
- Interacts with phosphate and deoxyribose with SSB
  - Electrostatic interactions - with phosphate
  - H-bonding to sugar
  - H bonds to DNA
  - Pi stacking - with Trp and Tyr
  - SEQUENCE INDEPENDENT



Lecture 20

April-23-17 4:46 PM

- Can proofread to ensure the correct nucleotide is added

2 DNA polymerases needed

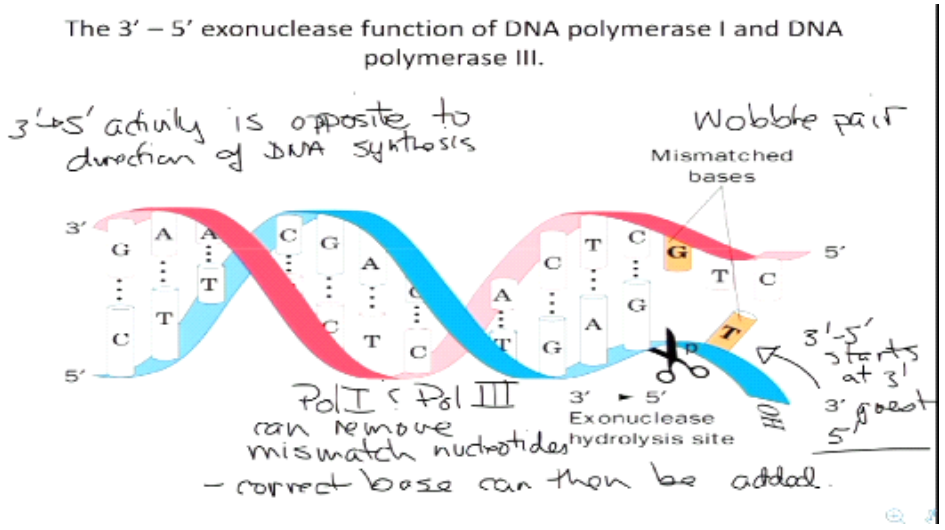
- *E. coli* has 2 DNA polymerases
- **DNA polymerase III (Pol III)** -- synthesizes the leading strand and most of the lagging strand.
  - very fast 1000 nt/sec
  - main enzyme for DNA synthesis
  - has proof reading activity ⇒ maintain fidelity of genome
- **DNA polymerase I (Pol I)** removes RNA primers and replaces them with DNA. This enzymes also has a 5' to 3' exonuclease activity.
  - discovered 1st (A. Kornberg 1956)
  - slow 10-20 nt/sec
  - can hydrolyze & replace RNA primer
  - works on lagging strand

- Both are high fidelity enzymes, RNAP has an error rate of 1/1000
- DNA replication error of 1/10^8
  - Low error rate maintains fidelity of genome
- Exonuclease happens at the ends, endo is in the middle

Proof reading function of PolI and PolIII

- RNAP error rate of 1 in 10^4 base pairs in *E. coli*.
- DNA replication error rate of 1 in 10^8 base pairs
  - low error rate maintains fidelity of genome
  - 4.6×10^6 bp in E coli genome
  - no errors!
- Pol I and Pol III have **3' → 5' exonuclease** activities.
  - enables proof reading!
- Removes newly synthesized 3' nucleotide of a daughter strand if wrong nucleotide was added
  - exonuclease activity occurs for wobble pairs ⇒ non-Watson Crick base pairs.
- Other enzymes also detect and correct errors in DNA
  - errors from UV or chemical damage
  - many error repair mechanism

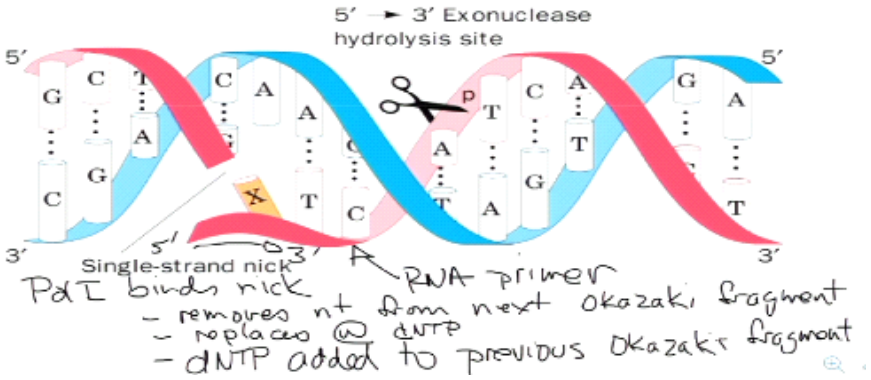
The 3' – 5' exonuclease function of DNA polymerase I and DNA polymerase III.



- Pol 1 uses 5'to 3' exonuclease activity to remove NT from next okazaki frament,

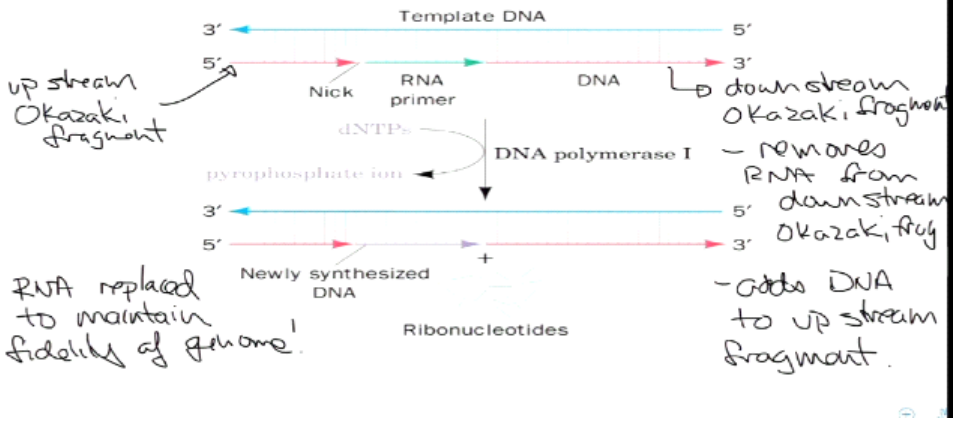
The 5′ – 3′ exonuclease function of DNA polymerase I.

RNA primer is replaced 1nt at a time



PolI binds nick
- removes nt from next okazaki fragment
- replaces w/ dNTP
- dNTP added to previous Okazaki fragment

Replacement of RNA primers by DNA in lagging strand synthesis.



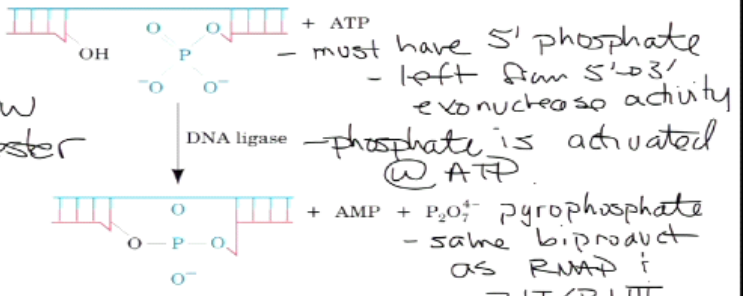
up stream Okazaki fragment

down stream Okazaki fragment
- removes RNA from downstream Okazaki frag
- adds DNA to up stream fragment

RNA replaced to maintain fidelity of genome!

Completion of Lagging strand synthesis

- nicks between discontinuous fragments are sealed by **DNA ligase.**
  - repairs phosphate backbone
- **Catalyzes the links of 3'-OH to 5'-phosphate groups.**



- must have 5' phosphate
- left from 5'→3' exonuclease activity
- phosphate is activated w/ ATP

makes new phosphodiester bond.

pyrophosphate
- same biproduct as RNAP i PolI/PolIII

Recombinant DNA technology

- DNA molecules formed by laboratory methods
  - combines DNA from many sources (recombines)
    - enabled by the highly conserved nature of nucleic acid biochem.
- Key tools:
  - Molecular cloning
    - cloning since the DNA is copied exactly
    - done in an organism by DNA replication
    - needs an extrachromosomal element
  - Polymerase chain reaction
    - enzymatic - in a test tube
    - no organism needed

Molecular cloning: the basics

- Isolate DNA to be cloned
  - usually a gene encoding a protein
- Digest DNA with restriction endonucleases
  - endonucleases cut dsDNA on both strands
  - very sequence selective! - if not selective an endonuclease would kill an organism by destroying its genome
  - require palindromic sequence
    5' CATATG 3'
    3' GTATAC 5'
    both 5'→3' sequences are the same

Molecular cloning: the basics

- Isolate DNA to be cloned
- Digest DNA with restriction endonucleases
- Ligate digested DNA into an extrachromosomal element
  - sticky ends overlap
  - ligate - use ATP to connect strands
    - like in DNA replication
    - use a ligase.
  - extrachromosomal element ⇒ plasmid
    - enable DNA to replicate in a cell.

Plasmid-based cloning vectors

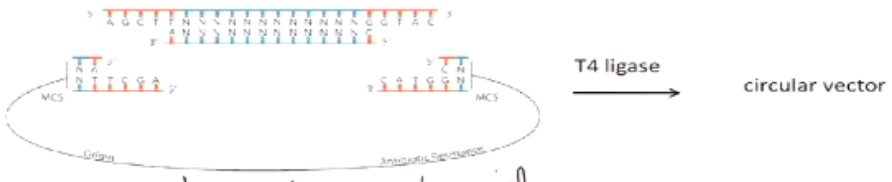
- **Plasmids** - circular DNA duplexes of 1 to 200 kb
  - contain the requisite genetic machinery (replication origin) necessary for autonomous replication
  - usually has antibiotic resistance gene (selectable marker)
- Types of plasmids are determined by their copy number.

- **Stringent control**- one to a few copies per cell.
  - low copy number
- **Relaxed control**- 10 to 700 copies per cell
  - high copy number
  - great for making lots of DNA!

Cloning of a gene into a plasmid

YFG – your favorite gene

gene of interest

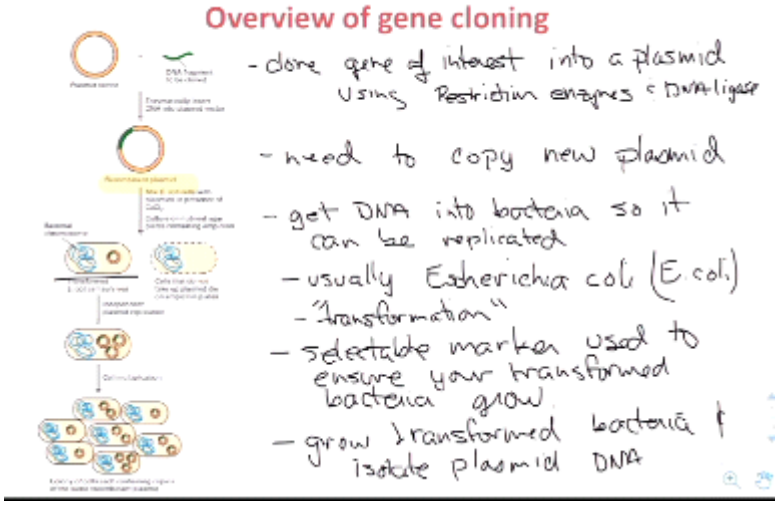


linearized plasmid (often called vector)

- only 1 way for the gene and plasmid to assemble due to sticky ends.

Lecture 21

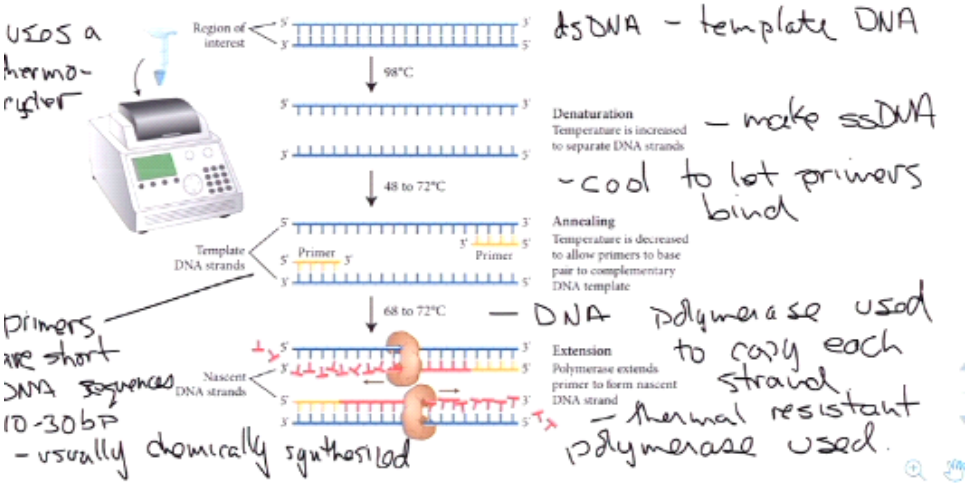
April-23-17 6:06 PM

- 

Invented in 1985 Kerry Mullis - Dancing naked in the mind field.

Polymerase Chain Reaction (PCR)

- no organism
- uses a thermo-cycler



dsDNA - template DNA

- make ssDNA
- cool to let primers bind
- DNA polymerase used to copy each strand
- thermal resistant polymerase used.

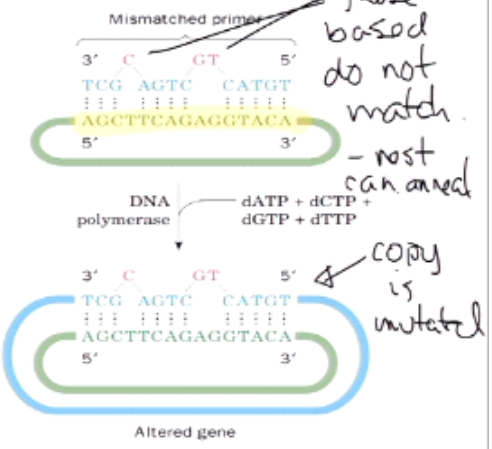
primers are short DNA sequences 10-30bp
- usually chemically synthesized

PCR can be used to create a mutation

Mismatched primer is used on a plasmid

Primer has altered sequence

- Amplification with DNA polymerase generates the mutated plasmid



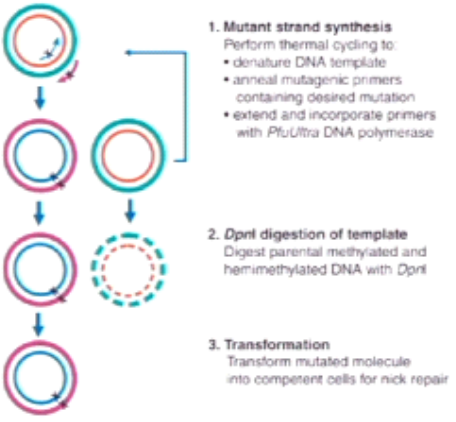
these based do not match. - most can anneal

copy is mutated

- can be used to change a specific codon! changing the amino acid in an encoded protein
↳ "site directed mutagenesis"

Template DNA must be removed

- After PCR wild type (wt) template still remains
- must remove wt plasmid.
- plasmid template usually isolated from bacterial
  - DNA is methylated in bacterial
  - PCR does not make methylated DNA ⇒ easy to distinguish!



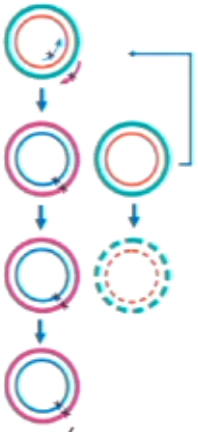
- DNA adenine methyltransferase
- Occurs at GTAC site - adenine methylation
- Does it affect major or minor groove?

Template DNA must be removed

DpnI is selective for
me
- GATC -
- CTAG -
me

- only digests wt template!
- PCR product is not methylated ∴ not digested

How frequent are GATC sites $1/4 \times 1/4 \times 4 \times 4 = 1/256$



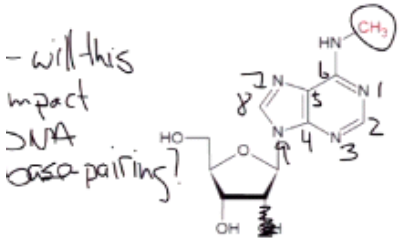
1. **Mutant strand synthesis**
Perform thermal cycling to:
• denature DNA template
• anneal mutagenic primers containing desired mutation
• extend and incorporate primers with *PfuUltra* DNA polymerase

2. ***DpnI* digestion of template**
Digest parental methylated and hemimethylated DNA with *DpnI*

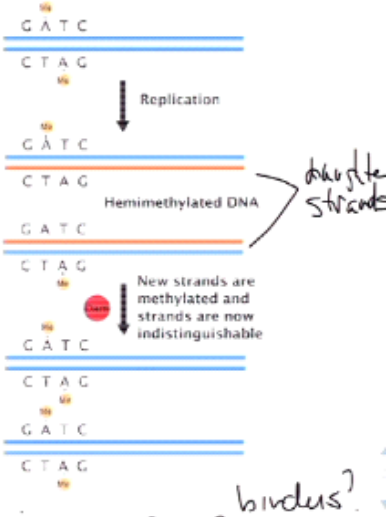
3. **Transformation**
Transform mutated molecule into competent cells for nick repair

DNA Adenine methylation (Dam)

A methylated at N6



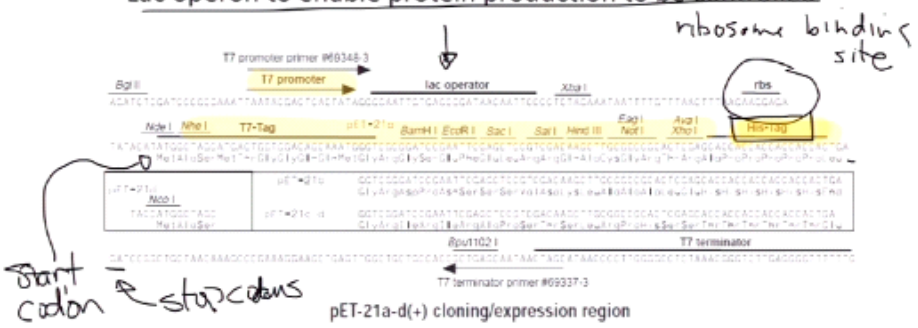
- will this impact DNA base pairing?



- bacteria add Methyl to protect genomic DNA from restriction enzymes.
- can also impact transcription factor binding to DNA
  - major or minor groove binders?

Inducible recombinant protein production

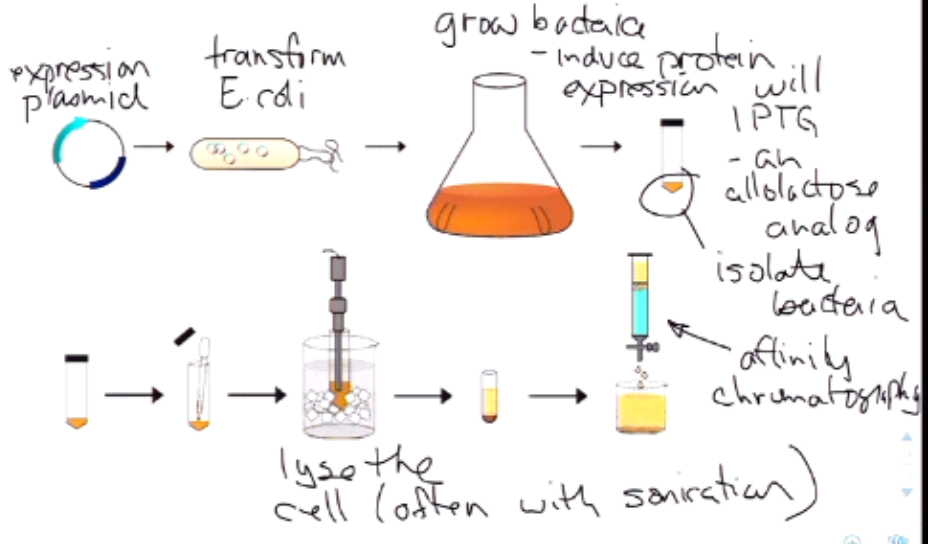
- T7 promoter from the T7 phage
  - highly active promoter 5x > than E coli RNAP
- Lac operon to enable protein production to be controlled



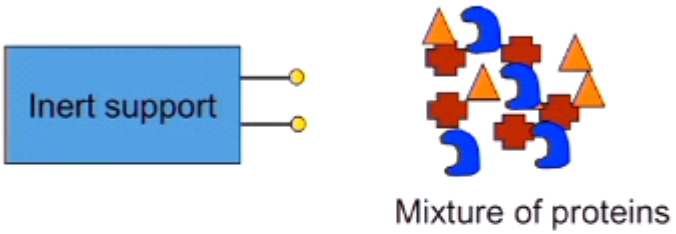
pET-21a-d(+) cloning/expression region

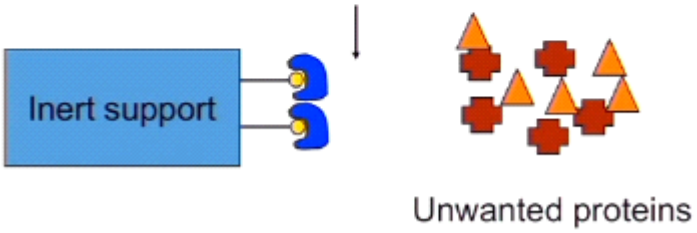
- Adds a C-terminal His tag for protein purification by IMAC
  - makes purification of recombinant protein easy!
- Affinity chromatography - isolate using His tag

Protein produced in bacteria by transcription and translation



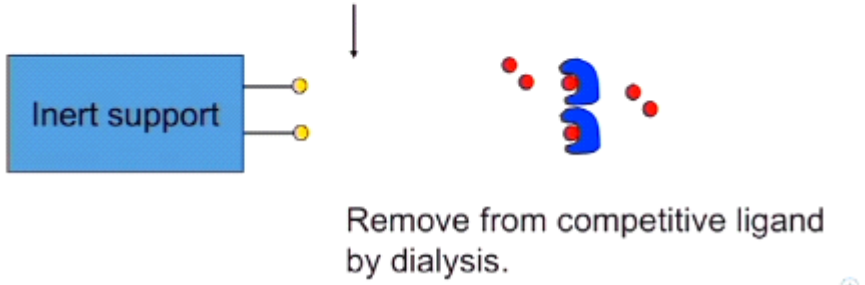
Affinity chromatography





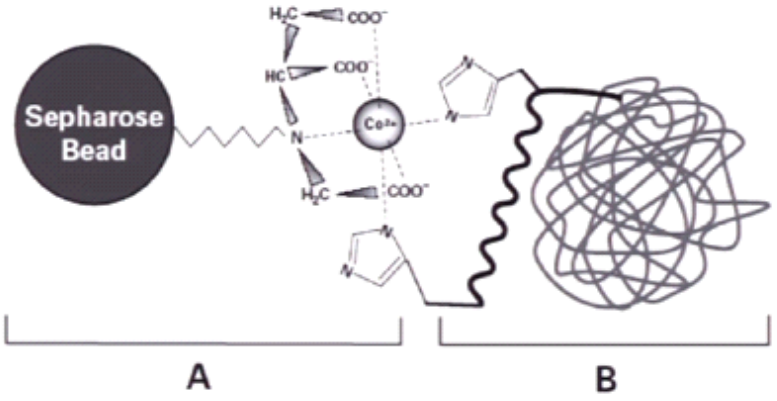
- Uses an affinity tag that will bind the protein of interest
- Elute with a ligand to remove the protein off the support

Affinity chromatography



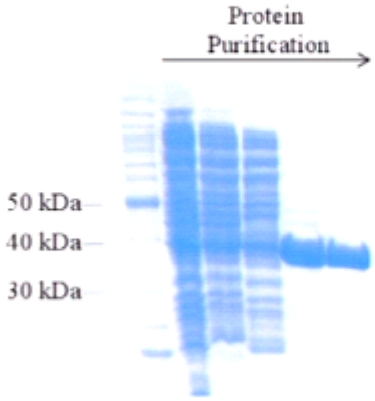
- Histidine binds to metal ions - ligands
  - Histidine binds to the Cobalt ion

Metal affinity chromatography (MAC)



- hex His taggod proteins bind TO resin
- only protein @ hexa his tag is your recombinant protein

Metal affinity chromatography (MAC)



- very effective at generating high purity recombinant protein!