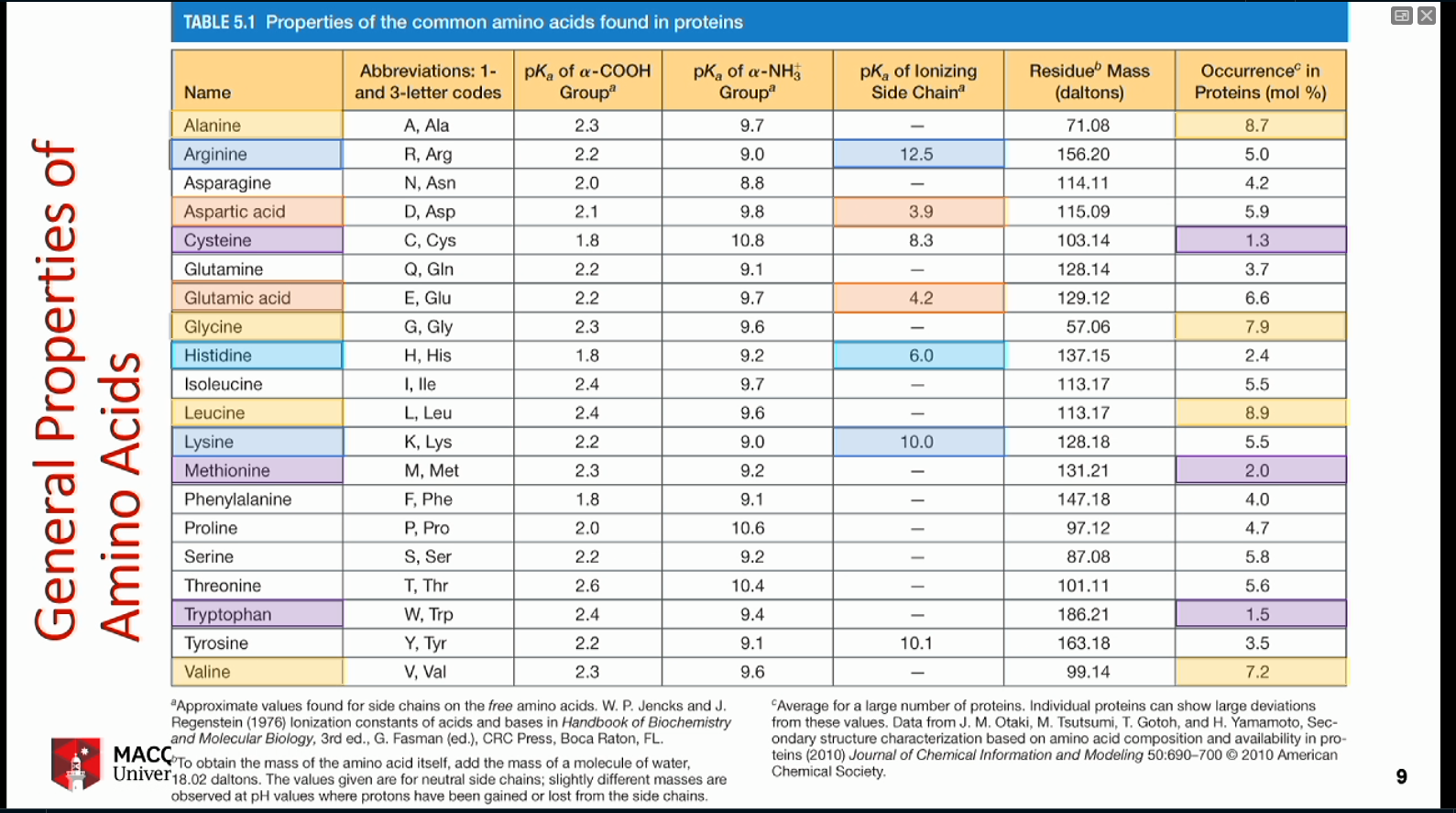
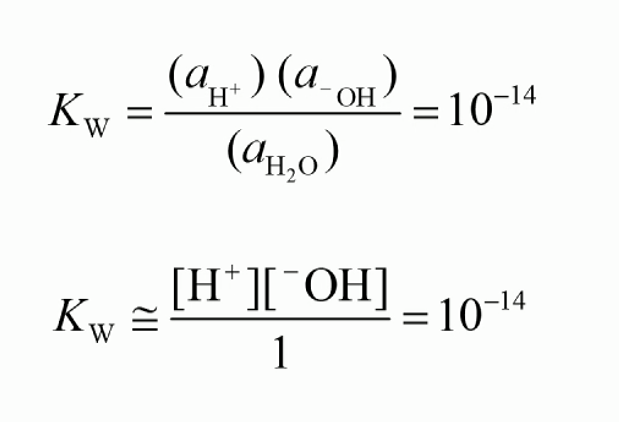
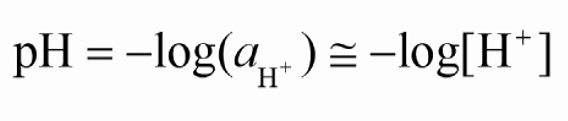
* Trace elements are 4th cell
* Some did create amino acids; some created nucleic acid but no living things
* Ancient cells had RNA
* Condensation is the opposite of hydrolysis
* Seclusion – breaking and making things at the same time and different places – keep them independent – no wasting
* Many molecules associate without non-covalent bonds
* Dimerise?
* Growth hormone stimulates cell growth and differentiation – without forming chemical bonds
* Hormone is protein – receptor is also protein
* Among all compounds that have H; HX – Only water is liquid – unique – the shape of water and the two lone pairs – very electronegative – like to form bonds with other hydrocarbons – on one side wants to donate – another side wants to gain; liquid molecules that require highest energy to heat up by 1 degree – requires a lot of energy to boil water
* Lipids are hydrophobic – hate water – clathrate – cage structure – nonpolar surface areas get surrounded
* 1 armstrong is the length of hydrogen
* Amino acids are building blocks of proteins
* Each amino acid has a common part
  + Alpha amino group that has alpha carbon
  + Alpha carboxylic acid – acidic end
  + Amino group is the organic basic group
  + Form nh3+ in water – carboxylic group will become charged – COO-
  + Side chain is different from one amino acid to another
  + Chirality
  + Optically active asymmetric chiral molecule
  + At neutral pH, as it has both positive and negative charge the net charge is 0 -> zwitterion
* Carbon is usually shown as black or green
* Oxygen is usually red – cuz it represents blood, life – haemoglobin
* Nitrogen is blue – also called azote – thunder has lots of charges and thunder is blue – nitrogen becomes ammonia
* All amino acid except glycine has asymmetric alpha-carbon
* Glycine has 2 H so not asymmetric
* D-amino acid can make immune responses
* There are more than 20 amino acids but only focus on 20 amino acids cuz humans
* Lots of amino acids are nonpolar (10 amino acids) - hydrophobic
* Methionine is the start codon of most genes
* 3 amino acid codes
* Rings make it hard to twist around – more stable
* 10 amino acids are polar - hydrophilic
* Polars do not lose protons – they form hydrogen bonds
* ¼ of the amino acids are charged – asparctic acid and glutamic acid are highly negative
* Glutamic acid in Chinese sauces – soy sauce, oyster sauce
* Histidine is basic but operate around very neutral pH
* Hydrophilic stays at the surface of the protein while hydrophobics are wrapped inside
* Blue is basic; red for acid; Histidine is mild basic

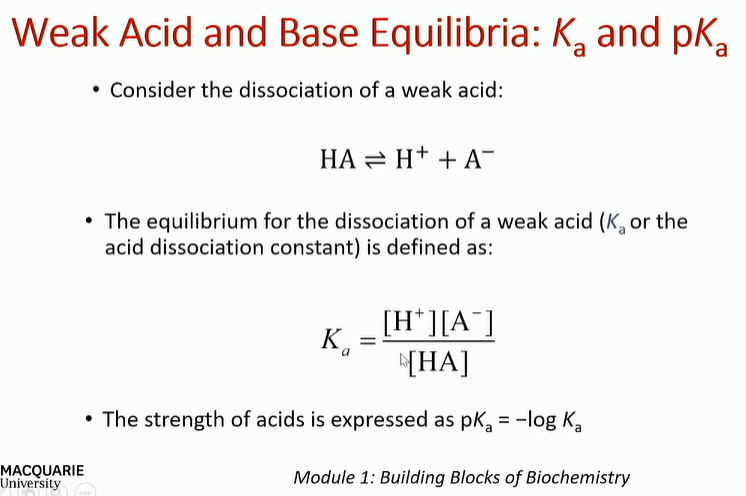


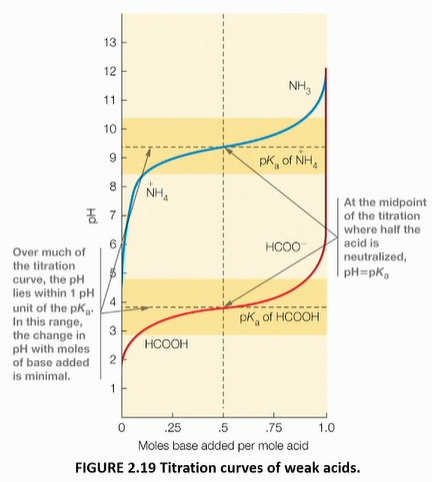
* Amino acids range in codon frequency – glycine, leucine, and valine are over-represent – hydrophobic – nonpolar – we like them a lot
* ATG has sulphur in side chain
* Tryptophan only has one codon – UGG
* Charged amino acid – equilibrium
* 0 to 14 is the Ph range
* Ammonium ion is NH4+ - can give away proton
* Imidazole is a mild base – has almost neutral Ph – the group found in Histidine
* 20 standard amino acids have common backbone but different side chains – only 5 are ionised (charged) that can interact
* Bases are proton acceptors
* Bronsted-Lowry definition
* Strong base associates completely into conjugate acid – organised – accept protons – rarely found in neutral form
* Strong acids will fully ionise
* Water is the principle source of proton to make acids charged – to accept protons from bases
* Water can self protonate
* Equilibrium means it goes back and forth – not fully in one direction – depending on temperature
* Make life easier by writing H3O+ as H+
* 1 water molecule gives equal amount of H+ and OH-
* Concentration is in [ ]
* Equilibrium constant defines ratio of the equation:



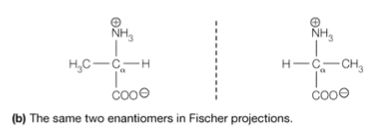
* Concentrationof water equals 1
* Kw = 10^-14 -> Proton = 10^7
* Water makes acid behaves like acid and base behaves like base
* Ion product is Kw (10-14)
* pH was proposed – potential/power/wtv of hydrogen – use log scale



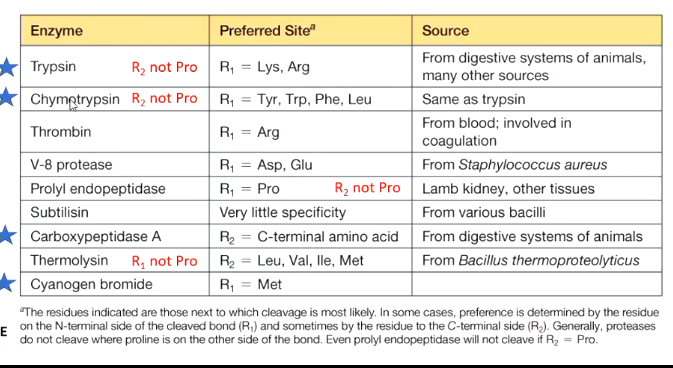
* log1 = 0
* 1 M HCl = 0; very acidic; occur in tummy but slightly less than 1 M
* 1 M NaOH = 14; very basic
* Most biochemistry occurs in pH in 6 to 8
* Bio reactions take place between pH 6.5 and 8
* Human operates in pH 7.4
* Water hardly is at pH 7; either basic or acidic
* Surface of proteins have polar charges – 50% are charged (like water)
  + Highly positively charged in low pH (acidic)
  + The same enzyme can react differently in different environment
  + When you get sick, your blood pH changes – enzyme changes – feel weak bcuz body is using a lot of energy
* Isoelectric point (pl) = 6.8
* 
* Ka is low bcuz we are dealing with weak acids and bases
* PKa measures the acidity of acid
* In chemistry, equilibrium is measured at 0 K
* In biochemistry, pKa is measured at 25 degrees
* Bicarbonate ion controls our blood pH
* Ammonium is toxic – ammonia goes into urea (non-toxic)
* Henderson-Hasselbalch
  + If ph = pka ; [HA] = [A-]
* Buffer range is the range of pH that the molecules do not change rapidly from base to acid and vice versa

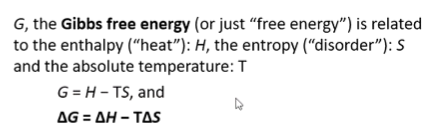
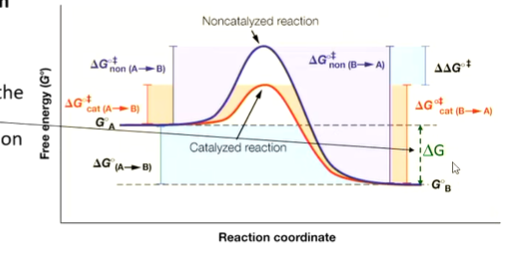
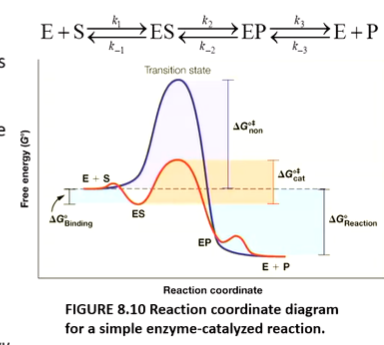


* Buffer will resist pH change within 1 unit of pH or 100 times of Ka
* Absorption of light – identifying and quantifying
* Modification of side chains
* Biological function of amino acids – some do things by themselves, some need to be modified
* Genes – how mRNA is made and transcribed by proteins for proteins
* Except glycine, has asymmetric carbon – C alpha is conncected to 4 H group – backbone has organic group, NH3, H, COO-
* Chirality is when side is not H, no double H
* Silica crystals grow in desserts, grow into spirals
* DNA also has left and right spirals
* L-Alanine – smallest amino acid – CH3 as side chain
* Fischer projections

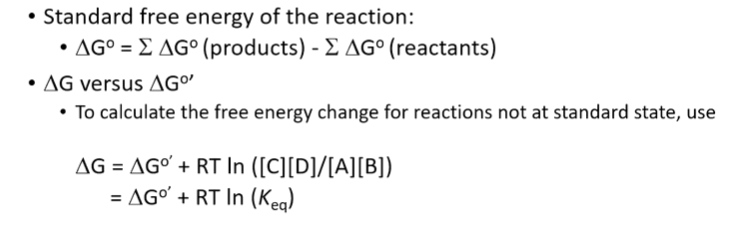
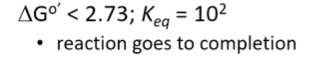
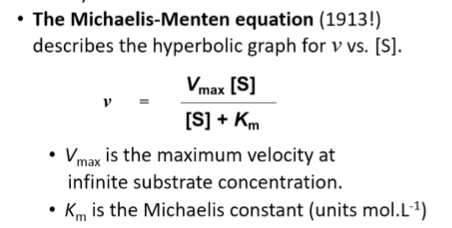
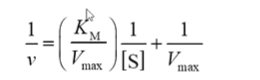
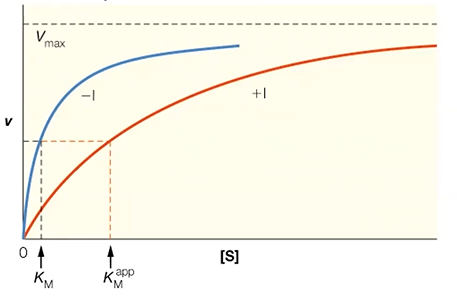
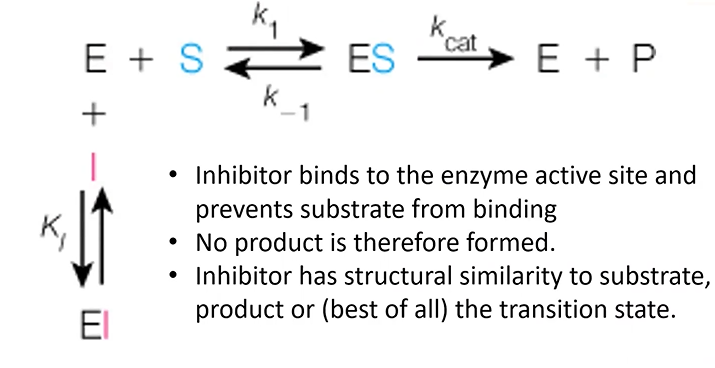
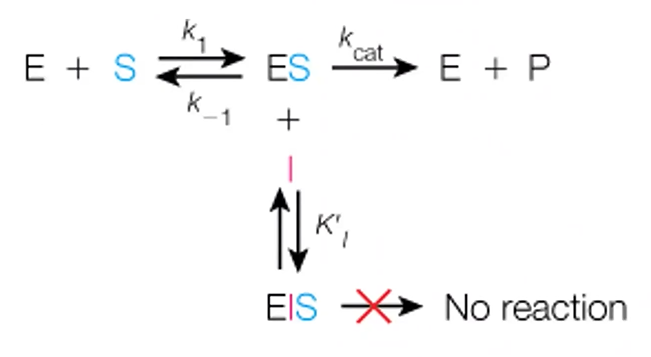
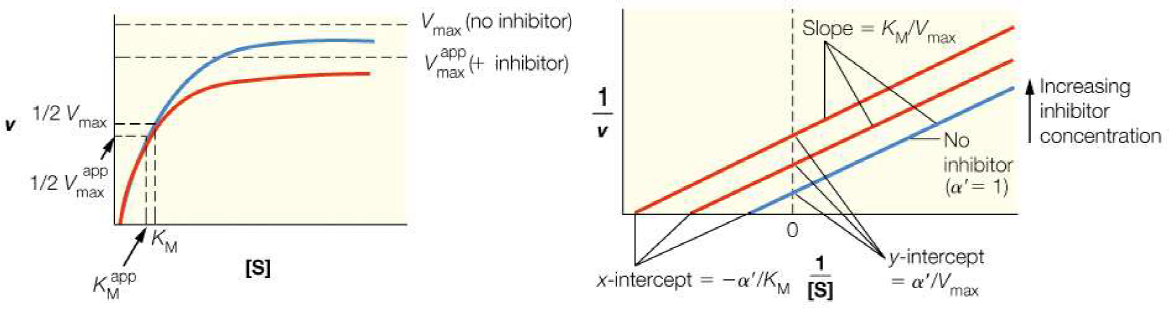
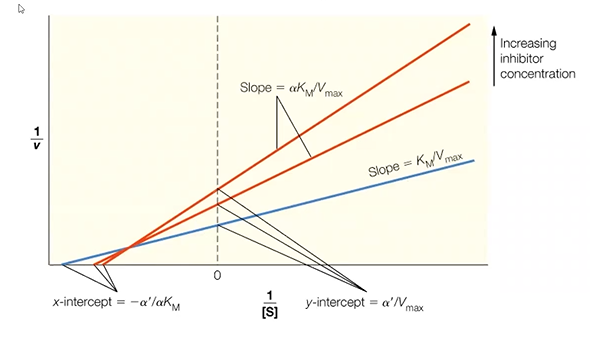


* L represents laevo – left hand
* D represents dextro – right hand
* One direction is +, one is –
* Twisting to left (polarisation) is L-molecule
* R stands for rectus (right)
* S stands for sinister (left)
* Humans have all L-amino acids
* L-Glyceraldehyde – smallest sugar
* Almost all sugars we encounter are Ds
* Configurations need to be remembered
* Names need to be remembered
* 50% of L and 50% of D when making in flasks
* Ibuprofen – only D-isomer has greater activity
* Biochemical reactions result in 100% of one stereo isomer – biotechnology saves energy and makes a lot more isomers
* Amino acids are colourless but they do absorb light in the near UV – in the particular range of frequency related to its energy/wavelength
* Nanometre – 10^-9 metres – proteins absorb in UV because of aromatic side chain but absorbance is fairly weak – peaks around 280 nm – a point where we expect accurate measurement
* DNA absorbs at 260 nm – strong absorbing
* We are using spectrophotometer – generate light and have specific filter to give light at specific wavelength – sample has to be in glass that has walls that are transparent – cuvette
* Molecules are absorbing light – we see them absorbing energy – there is a lost of energy that comes out
* Instrument measures light going out and light absorbed from the other side – give us the ratio
* Identification and …… are still done via spectrophotometer
* Concentration is measured in molar/mol/L
* Log of the ratio of the amount of light that goes in and light coming out – relevant to the distance that the light travels – make the distance l as small as possible
* C is the concentration
* Absorbance A = epsilon\*l\*c
* But l = 1
* Beer-Lambert’s Law – only works with dilute concentration – work with small concentration well
* This law can identify amino acids vs proteins vs nucleotides
* Side chains frequently get modified after the proteins are made – chemically modified by …………………..
* They can take very large molecules such as sugars
* Removal of phosphate deactivates proteins
* Hydroxyproline – important for nails
* Modifications are important for signalling pathways, stabilizing structures, gene expression/suppression…
* Sometimes are phosphorylated, sometimes are not
* N- and C- termini of amino acids are chemically active – form NH3 or COO- group – they can be modified or blocked (rendered inactive) – no chemical reaction can occur
* No chlorine in biochemistry
* Popular modification of cysteine residue – 2 armstrongs long bonds – longest bonds – 3D – stable structure – can occur across different chains
* Amino acids are biologically active
* Glycine and glutamate are neurotrasmitters
* …………………………
* Histamine is the compound that causes reaction – come from proline amino acid
* Dopamine comes from tyrosine
* Thyroxine comes from different groups – is the thyroid hormones
* Most important function of amino acids – their ability to link together
* Peptide bond – multiple amino acids joining together
* Everytime peptide bond occurs – lose water – called condensation
* The ends still have charges but the intermediate charges disappear – except side chains
* Reactions of proteins don’t occur spontaneously
* Protein synthesis occurs at the ribosome – need energy by hydrolysing ATP
* Only part at C-terminus can react – similar to 3’ end of DNA
* Residues – amino acids that are parts of proteins
* Side chains carry their chemical properties into the peptides/proteins
* Protein backbone forms on a plane – cannot weaver like a ribbons
* Peptides and proteins have lots of charges – side chains continue to have charges
* N and C terminus and 5 charged residues
* Each side chain can be titrated – has different pH – overall have pI
* pI changes every time depending on the order of amino acids
* DNA is transcribed to RNA
* mRNA is transcribed by ribosome
* proteins are folded up to 3D structures
* the 3D structures define the functions of proteins
* Up to 20 amino acids – called peptide – also called oligopeptide
* Many but don’t know how many – polypeptide
* Proteins normally have more than 40 aa
* Can be more than 1 protein chains – can form multimeric proteins – no bonds between the chains – noncovalent bond interaction
* Haemoglobins – heterotetramer – 2 alpha and 2 beta chains
* Proteins are polypeptides with defined sequences
* The numbers and orders of amino acids are restricted – can be mild variation – eg. human and whale myoglobin
* 92% are identical – quite similar in mammals but quite different in others
* Insulin – earliest protein to be sequenced – small enough to be fit on the slide – A chain and B chain are held together by disulphide bonds – independent disulphide bond to help the proteins fold
* Largest known protein is titin in muscle fibres
* Most are 100-1000 aa
* Many proteins have multiple subunits – sometimes bonded together like insulin but not always
* Genes to proteins
* Solving genetic code is very important
* U is RNA and replaces T
* AUG is the start codon – code for methionine – residue Met
* Some codes for multiple amino acids
* 4 DNA bases – purine and pyrimidines are major contributors to nucleic bases
* Pyrimidines – longer name but smaller structure
* Purines and pyrimidines are always attached to pentose sugars
* Nucleic acid has phosphorous on their backbones
* Adenosine is a hormone and neuromodulator – autacoid – sleep regulator
* When adenosine rises – we get up – caffeine works similarly
* Ribose is lower version of sugar – ribose has 5 Cs – sugar has 6
* Sequence is TGCA
* A pairs with T, C pairs G
* DNA transcribed to RNA, translated to proteins
* Each codon is made up of 3 bases……………………………………….
* Translation is done only in one direction
* Negative strand is used as the template
* Junk DNA – now noncoding regions – have important information
* Met tells ribosomes where to start translating
* Insulin – when they try to find protein strains, they found preproinsulin – very long chains
* Summary
  + Stereochemistry and light absorption are what we focus on
  + Amino acid derivatives help us perform different functions – hormones, activation, glycine and glutamate are actually neurotransmitters
* Pearson does not have back button
* Protein from dairy, non-dairy, meat, fish – 1/3 of diet consist of protein
* Protein performs all tasks in cells
* Enzymes are biological catalysts – very specific bio reaction – 10^16 times without enzymes
* Regulatory proteins – 14% - gate keeper – look after many proteins include enzymes – control which genes are turned on and off for transcription and activation
* Proteins generated in ribosome have to find their way outside of ribosome to perform their tasks – transport proteins carry proteins from one location to others
* Membrane transport proteins – facilitate transport ………….
* When supply falls short, storage protein comes into use
* Movement proteins – mobility eg. muscle cell
* Structural proteins – maintain cell wall, constitute skins etc. – very tough, cannot be digested
  + Collagen holds bones etc. together – when collagen hydrolyses, it becomes gelaton
* Signalling proteins – trigger actions for regulation to occur
  + Eg. scaffold or adapter proteins
  + Anchor proteins – hold them down – some proteins should not be moved
* Protective proteins – antibody, immunoglobins
* Defensive/exploitive – toxins and venoms
* Exotics – monellin – 2000 times sweeter than sugar
* pH and temperature affect protein stability – pH can cause proteins to aggregate or unfold (stop their functions)
* Assay – purify or quantify proteins
* Chemistry of side chains to maintain proteins and stabilise them
* Need to understand the protein we’re dealing with
* Presence of other enzymes, etc.
* A few degrees are enough to kill proteins
* Most times work in 0 degrees C
* Degradative enzymes – break cell wall and everything will spill out
* It’s because we are extracting proteins from cells – 0 degrees not 37 degrees
* Proteins are usually purified based on specific functions – solubility, ionic charge, mass, size, binding specificity
  + Will try to maintain optimum pH
  + Ionic charge – ion exchange chromatography is best
  + Affinity chromatography – not used in this unit
* Proteins are not fully soluble
  + Charge 0 is isoelectric point - pI
  + Solubility increases for small amount of addition of salts – salts make protective layers around proteins – salts charge so repel with proteins
  + Too much salt – super saturated – everything will settle down – salt prefers water so will remove water around proteins – proteins will find each other and stabilise
* Salting-out: Chuck in a lot of salt to get proteins out – ammonium sulphate is what we’ll use
* Electrophoresis depends on size and charge
  + Protein will be positively charged in acidic condition
  + Protons neutralise in alkaline/basic conditions – proteins will be negatively charged
  + Proteins are colourless
* PAGE uses gel filtration as well as movement – voltage will make charges moved – proteins at pI will not move – gel electrophoresis
* Ion exchange chromatography – the same as purifying water – hold back proteins that are oppositely charged
* Gel electrophoresis – about 3 hours
* Isoelectric focussing – separate charges – proteins move on gels bases on their pI value
* Isopycnic solution – vary from low to high density – spin the proteins – proteins will be settled based on their density – high density lower down
* Ultracentrifugation – tubes that can spin and then punctured – cannot reuse the tubes after punctured
* Chromatography – originally used to dye clothes – separate things based on colours – for amino acid mixture etc. – 4 to 5 hours to work – use porous substances – presence of proteins monitored by absportion spectrophotometry
* Column chromatography – add mixture to the top of the buffer – as mixture passes through – molecules fall down – molecules that don’t stick come down first
* Size exclusion chromatography – small gel beads that are porous – large molecules come out first – small ones get stuck in pores – can monitor appearance through chromatography
* Affinity chromatography – property of binding to another molecule – choose the molecule specifically – affinity matrix specifically chosen – His tag immobilises the proteins to the column – can later recover
* Purification – break the cells open – everything will spill out (proteins, cell membranes, etc) – separate soluble and insoluble – use affinity chromatography to remove the ones with His tag – apply ion chromatography – concentrate and purify proteins
* Column 5 of figure 5A – standards – compare mass of the proteins – is it right proteins or broken out into fragments – add the sequences to get the mass (eg. C, N, O)
* Dalton/Kilo Dalton is the mass
* Mass does not tell the function
* Whole genome sequencing
* Interested to find functions of proteins – use bioinformatics
* Compare human myoglobin to human alpha chain of haemoglobin – not 100% but quite similar – good match – probability of this match is very small – not by chance
* At least 25% similar – homologous
* If the side chains are similar – 37% like – can substitute each other
* Lots of gaps are not good – conserve functional parts but getting changed over time
* Best matches have very low scores
* Homology search can be done against multiple organisms to understand evolution
* Cytochrome c in animals and plants – very similar
* Help us understand mitochondrial respiration – organisms have highly similar sequences – highly conserved protein – highly conserved sequences
* If it’s absolutely conserved in the 27 sequences, there will be no change – these positions are very important – cannot mutate because will lose function
* Help us understand diseases such as cystic fibrosis (first to be found as genetic disease) – caused by one single change in amino acid – the whole codon is missing – because of that, protein does not function
* Insulin has 2 N-termini and 2 C-termini – bind by disulphide bond
* Oxidise bond to break them
* Protect the disulphide position – cysteine residue – so that they don’t form disulphide bond and stick up again
* Sequence of entire insulin has 3 bonds
* Two chemical methods by which we can detect how many protein chains
  + Dinitrofluorobenzene – DNFB – add DNFB to get polypeptide (shorten it)
  + The aromatic ring can be cleaved out later to get the amino acid
  + Dansyl chloride – cleaves first aa off but the rest of polypeptide also chopped up into polypeptides – don’t get long chains – get free polypeptides
  + Pay attention to these methods
* Carboxyl terminal can be determined by using carboxypeptides – cleave off last residue – leaving new shortened amino acids
* Haemoglobin – both N-terminal of alpha and beta chains are the same – we need to know both chains
* Use beta-mercaptoethanol to break disulphide bond and then iodoacetate to keep them from sticking
* Use performic acid to cleave the bond in one step – form sulphate to prevent bonding
* After having protein chains, need to cut them into small sequences to sequence them
* Two different methods to assemble them back in unambiguous way
* Trypsin and chymotrypsin cleave proteins at different points
* Met-specific – use reagent cyanogen bromide (highly toxic) – only recognise methionine – cut there
* Trypsin cuts after lysine or arginine – only 2 aa that are positively charged in neutral Ph – cut after C=O and create O=C-O-
* Cannot cut residue of proline
* The star will be tested
* R1 is the side chain before peptide (C-N) bond – R2 is after
* R2 forms the new N-terminal
* R1 after cut will be the new C-terminal
* Which enzyme cuts after which site



* Edman used phenylisothiocyanate (PITC) to cleave off the first aa – aa is removed and identified – the first polypeptide chain is sent back and treated with PITC – recycled many times – one at a time – 1st aa, 2nd aa, etc.
* To determine exact sequences of aa in the peptide fregments – use MS-based proteomics – MS stands for mass spectrometry – give out mass values – how heavy is each fragment
* Look at mass values of each aa
* Digest protein with trypsin – cut them up – last fragine is usually argine and lysine – sent to mass spectrometry – come out as lines – sent to software computer to get the results – choices of what the sequences may look like
* Myoglobin contains only one chain
* Peptides still contain bonds that can be broken up
* Some peptides are sent for subsequence breakup
* Can locate disulphide bond using MS
* Real MS looks very messy – we zoom in to find bits of interest to us – fragment it futher to get subset
* Individual MS can be sent for second time to get more detailed information
* Fit peptide sequences together – chop in different places – look for regions of overlaps – how to put them together to get our desired sequence – help from softwares
* Overlap allows determination of aa sequence without doubt – use chemical reagent such as CNBr that looks for Met – trysin that cuts after arg and lys – matching of peptides using overlapping sequences give the long aa sequences
* GenBank is largest DNA bank
* Protein sequencing helps us understand what proteins do – large scale sequencing shows a lot of aa repeats – gene duplication
* Domains – similarly made proteins – have independent identity – in combination provide functionality
* A lot domains came from stealing from other organisms or common ancestors
* Domains mean parts of proteins that are also seen in other proteins – common bits of proteins that are seen in different proteins
* Try to summarise steps in sequencing proteins – why important to identify N-teminal – how? – explain why long polypeptides have to be broken up to small peptide fragments for sequencing – Prac 3 might be useful
* Protein structure – multiple levels
  + 3D is functional – form 3D shape that provides structure – come from small bits of local structures
  + Secondary structure – held by backbone hydrogen bonds
  + Quaternary structure – several subunits assemble – already folded up but assemble through non-covalent interaction – form functional units – last level of protein structure
* Haemoglobin
  + Primary structure is the aa sequence – 2 different chains – alpha and beta chains
  + Some parts of chains adopt repeatitive seq
  + The backbone curls up into spirals or springs – alphahelix – in globin (myoglobin etc) was predicted even before tested – responsible by hydrogen bond – got the name alpha because it was the first one
  + Helices are very compact – provide flexibility to structure
  + Tertiary structure – myoglobin? – helices assembled into a shape that facilitates binding of iron carrier group called haem – bound by certain aa to proteins – middle of flat assembly is the iron atom that holds oxygen
  + Quaternary – 2 alpha and 2 beta subunits
* Peptide bond is very unusual – neither single nor double – C-N is 1.49 A – C=N is 1.27 but peptide is 1.32 A
  + Partially double bond – because oxygen has 2 lone pairs but H has 1 lone pair not as strong as O at holding back electrons – lone pairs sometimes is between C and H – Oxygen binds with N – sp2
  + Usually see trans peptide bond – cis is very rare – Van de wal repulsion
* Rotation about alpha-C is the only free rotation
  + phi plane and psi plane
  + steric clash – when conformation of phi = psi = 0 – not allowed in polypeptides
  + naturally occurring secondary structures – no steric clash
  + right-handed alpha helix – left-handed is very rare in proteins
  + fully extended polypeptide chain – nothing is hitting anything else – beta sheet – 2 rows that can hydrogen bond each other – backbone also held by H bond
  + each sheet is composed of a single element called a strand – segment of polypeptide chains that is stretched to its max – fully extended
  + Proline – in such as collagen – structure protein – 310 helix forms forms tight spiral – mainly proline
  + Alpha – starting point is aa end – free N-terminus that has lone pairs – slightly positively charged – top has C group – negative ends – automatically dipolar – positive attracts to end of helix
* Peptide bond imposes planarity on the 6 backbone atoms of each peptide – free rotation only allowed in alpha C – dictate what types of assembly are possible – after proteins come out from ribosomes – automatically find a stable structure – without any external interaction - provide limited choices of how these sequences can fold up – both alpha and beta structures are stabilised by backbone H bond (non-covalent bond)
* Secondary structure – only about 80% of proteins are folded in regular ways – the rest are unstructured
* Fibrous proteins – special proteins that cannot be degraded by trpsins and chemotrypsin – cardilage and muscle tissues
* Globular proteins – spell up into specific shape – biological functions – tertiary structure – the biological function is connected with shapes
* Dynamic of globular protein structure – certain flexibility to bind things
* Prediction – amino acid sequence – use computer to predict 2nd and 3rd structures
* Local folding is secondary
* Naturally occurring secondary structure – right handed alpha helix, beta sheet, 310 helix – backbone held by hydrogen bond – thousand bonds in DNA – H bond on its own is weak
* Alpha sheet - Yellow lines are H bond – side chains stick out – backbone like a circle (wheel etc) – using 3D viewer can rotate proteins
* Beta - Secondary structure – backbone like a plain – sidechains going above and below plain – hydrophobic may be on a face
* Only 2 types of rotation are permitted – planarity enforces that the backbone between one and another C-alpha are all in one plane – forces proteins to only adopt certain changes
* Steric interaction determines how proteins are folded – only certain foldings are allowed
* Alpha helix – closes packing of backbone and atoms – most common in protein structures
* Simple poly-L-alanine – only certain angles are permitted - +180 and -180 are converts
* Ramachandran plot – Ramachandran is the scientist name who figures out those angles
* Fibrous proteins – a lot of glycines, alanine, serine, cystine – collagen has a lot of proline – everything else very low
* Specific structural proteins – keratin, fibroin, collagen – very peculiar – not providing lots of arginine and aromatic compounds – cannot be chewed down by trypsin – unlike normal proteins – only have few proteins – which predetermine what structures they form
* Keratin forms alpha helix – coils region – form long fibres and twist around each other – twisting provides strength
* Silk – softness and flexibility but extremely strong – beta sheet – strength is attributed to the layers that can sit on top of each other and difficult to break apart
* Collagen – connective tissues – a lot of proline – hydroxyl proline and hydroxyl lysine – can assemble as multi-stranded helices – vitamin C is important for proline hydroxylation – vitamin C deficiency causes scurvy
* Collagen – 3 strands bind together – form strong pattern – lots of diseases are caused by mutagen to collagen sequences during protein making
* Alpha helix - Hydrogen bonding between first and fifth residues – second to sixth – i+4 – first carbonyl group bonded to fifth NH – oxygen has negative dipole – H has positive dipole – create vertical ladder – build up H bonding pattern – end of helix has free NH that is slightly positive charge – other end has free CO which is slightly negative
* Proline doesn’t have H on N because side chain comes back to bind N – proline not happy with alpha-helix – occur before or at the end of proline
* Beta-sheet – H bond between the strand – antiparallel – almost 180 degrees in H bond – most stable
* Parallel sheet – crooked H bond – H bond not 180 maybe 160 – 170 – not as stable
* The chains have to be connected – the loop – can be U bend or go back and form parallel beta sheet – if they turn around called turns
* Turns can be stabilised by H bond – glycine loves to be in turns – cuz no side chains – no steric
* Loops provide stability
* Chou table – may appear in exam
* 2nd structure compacting themselves form 3D
* Globular proteins form 3D – found globules in proteins – call them globular proteins
* There can be mixtures of helices and strands – cartoon can show spirals and strands – 2nd structure hidden from view but does exist
* Stands are like arrows – arrows can tell parallel or antiparallel
* 6.15a – mostly antiparallels – but also have 1 parallel
* Stick models show atoms – H bond between CO and NH – H bond dotted lines – couldn’t see H cause too tiny
* Surface models – can be coloured in different ways to tell us what aa at the surface – atom colouring not so helpful – positive (blue) and negative (red) more helpful
* MSEP – calculated view
* Experimentally – proteins are salts
* Electron microscope of collagen – big ones use this one – more reliable
* X-ray diffraction – better zoom
* NMR for soluble proteins – show that proteins can flex – get multiple structures from NMR
* Protein Data Bank – all protein structures – only one unlike sequences
* X-ray diffraction – X-ray getting reflected by nuclei – can get pattern of proteins – patterns are then analysed – using computers now – originally by hands and maths – computer program looks for blop of density and fit side chains into them – sometimes don’t see side chains but imagine that it is there
* Crystal diffraction – collect images at different resolution 4 armstrongs more detailed than 16
* Most crystallography collected at 2 armstrongs – coming out at 1.3-1.4 – but still can’t figure out density of H – H is too small – which is why we don’t show H in crystal structures
* NMR – shows proton shifts so can see hydrogen bond location – have to be small and soluble proteins – spectrum that shows proton shifts – water (solvent) has a single shift
* Without the folded shape, proteins cannot have proper structures – different architecture – may contain many domains – one protein can have 2 or 3 folds
* Typical domain has 200 aa
* Domain retains its own function
* At least 40 enzymes that bind the same co-factor – structure passes on to different enzyme
* Classification of protein structure based on helices and sheets
* Turns are short, regularly structured regions that connect 2nd structures
* Parallel beta – call alpha + beta
* Few 2nd structures – call disorder proteins – regular proteins – expressed and translate – have function but don’t have shapes
* Common features – hydrophilic exterior and hydrophobic interior
* H bond in the backbone is the main stabilising structure
* Beta sheets twisted and wrapped into barrel
* Polypeptide chain can turn corners – beta or gamma turns
* When look at sequence, hydrophobic residues are green, hydrophilics are pink, not sure in black – does not make sense in writing sequences but make sense when fold into 3D
* Hydrophobicity is the main driving force to form 3D structure
* Nonpolar residues prefer hydrophobic interior
* Charged polar residues contact with aqueous
* Uncharged prefer surface but when buried, can form H bond
* Thermodynamics of how proteins fold – just know that there is a driving force
* H bonds are typically noncovalent interaction
* Disulphides are covalent interaction
* Disulphide bonds are main contributors in small proteins
* We see that different proteins adopted different shapes and methodology to stabilise themselves
* Zinc finger binds DNA – have zinc ion that hold the protein in shape – start transcription
* Heme group makes myoglobin compact and red
* Proteins need to be maintained at particular temperature, pH, detergents, etc
* Denaturation is when proteins lose their shapes – lose functions too
* pH maintains shape and charged – H bond
* Detergents can invert the proteins – stabilise the hydrophobic side chains – ruin the shapes
* Guanidium ion – at the end of arginine – is also an independent ion – and urea – N compounds can cause proteins to unravel – called chaotropic agents
* Ribonuclease A – Add urea – break disulphide bond – become cystine – proteins just swim around – can also oxidise the denatured ribonuclease A with BME
* The funnel – thermodynamic folding – the lowest energy part is the native state at the bottom – along the rim is the denatured proteins – some structures lead down into the hole – some lead to local cavity – so large proteins can misfold – protease comes to cut them out
* Chaperonins hold large proteins – let them fold and unfold properly
* Misfoldings lead to diseases
* Slide 45 – dimer of molecules called transthyretin – one yellow – one magneta – show assembly that there is a centre of symmetry shown by black dots – two molecules appear to be twisted around each other – sometimes have hydrogen bond – 2-fold dimer
* Cartoon – no side chains shown
* Tetrameric enayme phosphor… - 4 subunits – one centre of symmetry – ending in –ase meaning enzyme
* Does have some math – ignore equations – except the one she mentions to remember – no need to derive anything
* Co-factor helps them to perform their tasks efficiently
* Energy term
* Free energy – thermodynamic quantity – driving force for enzyme direction – have limitation – clever way to overcome unfavourable energetics
* Enzyme kinetics – how rates of enzyme can be determined
* Enzymes are proteins – 20% proteins are enzymes – catalysts are prime function – increase rate of reaction without being changed – can participate but has to be regenerated – accelerate equilibrium without changing thermodynamic
* Lower energy barrier – activation energy of transition state – the way enzyme works
* Called agents of metabolic change – coined in 1878
* Cannot catalyst reactions that are energetically unfavourable
* Catalyst that reduces activation energy
* Carbonic acid formation – carbon dissolves in water to cause H2CO3 – reaction super slow – 10 million times faster with enzyme
* Won’t be able to perform tasks without enzyme
* Speed up is normally million to 10^12 with enzymes
* The enzymes perform rate enhancement under mild reaction condition
* All reactions occure under 100 degrees at P = 1 atm, almost neutral pH
* Greate specificity in enzymes – both substrates and products
* Enzyme has to be controlled and regulated – allosteric or covalent modifications
* Controlled amount of enzyme by controlling rates of gene expression, transcription, translation
* 6 categories of enzymes
  + Most important is oxidoreductases – oxidation reduction
  + Transferases – transfer functional group - eg. kinase - transfer phosphate from ATP to another molecule
  + Hydrolases – opposite of condensation – adding water to covalent bond
  + Lyases – group elimination to form double bonds
  + Isomerases – transform molecules so that they become substrates
  + Ligases – stick things together – done by using energy from ATP hydrolysis
* Name of enzymes – tell what the enzymes do - eg. glucose-6-phosphatase
  + Ending with “-ase” meaning enzyme
  + Use glulcose-6-phosphate as substrate
  + Phosphatase – remove phosphate – hydrolatic enzyme from the substrate group
* Major classes of enzymes examples
  + alcohol dehydrogenase – oxidoreductases – break down ethanol into acetaldehydes
  + hexokinase – transferase - add phosphate group to glucose for glucose breakdown
  + hydrolases - carbosypeptidase – cut down the last amino acid in peptide chain
  + ligases - pyruvate carboxylase – take pyruvate – add carboxyl group
* Enzymes are very selective – prefer specific substrates – specific stereochemistry – work very well for particular one or two substrates – stereospecificity
* Substrate specificity – chymotrypsin - break down protein by recognising aromatic residuals – will cut peptide bond immediately if the next one is not proline – if feed ester – breakdown but not efficient
* Some enzymes need cofactors – can be metal ions or coenzymes – coenzymes have cosubstrates (participate – act as substrate – has to be regenerated) and prosthetic groups (do not change – provide environment)
* The ions and coenzymes have 3D structures
* Vitamins are coenzymes – perform specific reactions
  + Nicotinamide adenine dinucleotide – vitamin B3 – called niacin in free form
  + NAD+ - has nicotinamide, D-Ribose, adenosine – oxidised form has benzene ring
  + NADP+
  + NAD is the participant when we metabolise ethanol to acetaldehyde by ADH – NAD+ becomes NADH and ethanol is oxidised
  + NADH is regenerated by other chemical processes and fed back to the reaction
  + Without NAD+, the enzyme will not function
* Metal ions
  + Ion and cytochrome oxidase
  + Zinc and alcohol dehydrogenase
  + Cobalt and cobalamin – vitamin B12
  + Eg. zinc helping carboxypeptidase A enzyme –carboxypeptidase A takes off the very last residue in a protein – zinc in present of solvent (water) able to hydrolyse the bond and free the N-terminal and COOH group
* Apo and holo enzymes
  + Apoenzyme – inactive without the bound cofactor
  + Holoenzyme – has cofactor bound – active
  + Complexes where substrate is bound – ternary complex
  + Coenzymes which can be chemically changed but need to be regenerated for the enzyme to continue function
* Thermodynamics:
  + Delta G
  + The rate of chemical reaction depends on many factors
  + To know whether the chemical reaction occurs or not – need to go back to thermodynamic
  + Free energy difference between the initial and final state – refer to products and substrates
  + Gibbs free energy – G
  + Delta G is the measure of the change from substrate to product
  + 
  + Delta G represents a portion of the energy that is required to do useful work at constant temperature and pressure
  + If delta G is negative – product is lower energy than reactant – thermodynamically favourable – exergonic process
  + If delta G is zero – reversible – no extra energy to do work – but can balance the budget – equilibrium process
  + Positive – product higher energy than reactants – difficult as need to pump up reactants to become products – unfavourable – endergonic process
* Transition state – everything is assembled – no reaction – no product – determine the rate of reaction – because the higher tge transition state in energy, the more difficult it is to climb the hill
  + The rate depends on order, concentration, temperature, rate constant
  + Free energy has to be negative
  + Initial and final state also go through transition state – additional energy barrier to convert reactants to products if the reaction is reversible
  + The transition state has more free energy – unstable – can switch and form either products or reactants depending on where it is headed
* The change from reactant to product represented on x-axis – fee energy on y
  + To form products – have to go through top of hill – transition state
  + Purple is noncatalysed reaction
  + Red reaction is what enzyme does – speed up transition from reactants to products
  + Enzymes bring transitional state to stability and bring it down
  + Once enzyme has brought down the hill, the system checks which side is favourable for spontaneity
  + When product is lower energy than reactant – negative delta G – favourable – will get products
  + Transition state does not go back to form the reactants but form the products
  + 
* Enzymes preferentially bind the transition state of the reaction
  + Not as simple as it seems
  + Number of adjustments that need to be made
  + Reactants bind
  + Enzyme-substrate complex (ES) is formed to hold the pieces of the puzzles
  + Transition state which transforms into products
  + Enzymes regenerated
  + Energetics is complicated
  + ES – small local minima
  + Es go through transition state
  + Break down into enzyme and product
  + A number of delta G values throughout the pathway – but we are interested in the delta G of the substrate and the product which drives the reaction
  + 



* + Transition state binds best to the enzyme has greatest affinity to the enzyme and complimentary in terms of shapes and electrostatics to the transition state
  + Transition state analogues are often used as inhibitors – extension putting methyl etc – inhibitors in biotech industry and drug designs
* Enzyme inhibitors are used as drugs
  + Real drug molecule
  + The three drugs on slides – Work on HIV protease
  + 1 prevent transcription
  + 2 prevent virus from replicating
  + Protease – 3 preventing it from breaking the cell wall
* Transition state analogues as enzyme inhibitors
  + L-proline forming D-proline in bacterium
  + Proline racemase is the enzyme
  + Need to go through planar transition state in order to change the configuration
  + Pyrrole… and delta… mimic the transition state and stop the bacteria from generating D-proline
* General chemical reaction
  + Most biochemical reactions are reversible reactions
  + Reactants – initial thermodynamics
  + A and B are reactants
  + C and D are products
  + Compare energy on the right side to the left
  + Apply constant temperature and pressure
  + Delta G = 0 -> at equilibrium
  + Balance between amount of products and reactants – constant concentration – equilibrium constant controls this
  + Delta G negative -> get products -> forward reaction occurs spontaneously
  + If positive -> reactants would accumulate -> no forward reaction
  + If delta g is large – only one direction predominates
  + Equilibrium constant is the product of concentration of products divided by reactants
  + Delta g for such reaction can be calculated if we know products and reactants: delta G = products – constants
  + 
  + ln is the natural log – not log base 10
  + Fine line of -11.2 Kj/mole – enough to push equilibrium to one direction
  + 
* Biological systems need energy to perform work –
  + Mechanical work - flagella rotation, muscular work, concentration
  + Concentration and electrical work – moving charges through, osmotic changes, etc
  + Synthetic work – changes to chemical bonds – how we cut food up and reassemble them
  + Energy comes from hydrolysis of high energy phosphate bonds (ATP, ADP) and reduced conenzymes (NADH etc)
* To perform energetically unfavourable work (positive delta G)
  + Add 1 phosphate group to glucose to make glucose-6-phosphate
  + Hydrolyse ATP - producing ADP and release a lot of energy
  + Overall energy delta G negative
  + That’s how we get energy from sugar
* Substrates can form complex with enzymes – ES complex breaks down – resulting in E and P
  + Energy of A and P are the same with and without enzymes
  + Enzymes cannot change the delta G but can change the height of the transition state
* Rate of reaction
  + How much P froms from S
  + Rate is represented as v
  + How much product we generate per second
  + V is not a constant number
  + V vs concentration:
    - With enzyme, get hyperbolic curve with a vmax
    - Without enzyme, get straight line
* Amount of substrate decreases and P increases as the reaction goes
* Steady state of ES
* E gets chewed up, regenerated and go back to the reaction
* We only need small quantity of E because it quickly forms ES
* E is usually not free – mostly ES
* V is usually when we have constant ES concentration
* Michaelis-Menten equation
* 
* When v – vmax/2, where it will be perpendicular to S, Km
* 1/v and 1/S – make the curve straight instead of hyperbolic
* 
* When km = [S]; v = vmax/2
* Units of Km is mol/L – inversely related to affinity of S for E
  + If small Km – need only a little susbstrate
* Catalic constant Kcat = vmax/[E]
* Kcat is the internal rate – how fast we can use this enzyme – measure the time a given site can turn S into P per second
* Kcat = vmax / [E]
* Kcat/Km measure enzyme efficiency – enzyme specificity constant – determine best S for E with low S
* Km is unique for each ES – low Km is good enzyme
* Large Kcat means large amount of P is being formed
* Inhibition and control are required to make enzymes work only when required
* The active site of protein – enzyme
* Active vs remote site
* Regulation – allosteric effectors or chemical modifications
* Active site is where one or more substrate molecules bind for the protein to carry out its function – usually comprises onl yfew residues from proteins – located far apart in aa sequence but close together when looking at 3D
* Mutating active side residue will completely destroy or modify E function
* Active site residues are localised to specific S binding pocket by folding of 3D structure – the entire folding process focuses on bringing the active site together for the chemical reaction to occure
* Complements the transition state in term of size, electrostatic, pH
* E binds both S and P but prefer the transition state – strongest binding affinity
* Serine proteases – 3 residues are critical for functions – histidine H57 H bonded to D102 and S195 – absolutely critical – ezymes in digestive system – trypsin, chymotrypsin and elastase – if stretch out will see 50 residues away between H57 and D102 and 100 residues for the 2 – but very close together when 3D- Histidine acts as base under physiological conditions
* Chymotrypsin cuts protein chain after aromatic residue – aromatic residue that cam fit into substrate specificity pocket
* Reaction mechanism won’t be tested – while the residues may form bonds with solvent or the substrate – the E will be regenerated and take another S
* Mutating any of the residue may destroy or decrease the functions of E
* The active site of E is a completely different environment ot the surrounding solution – it can be more acidic or basic by the positioning of the charged residues in the pocket – can provide hydrophobic environment – the S can only get into the pocket and fit into certain direction – make it stereospecific – only particular reactions can occur – electrophilic or nucleophlic functionality groups
* If we know the exact structure of the active site – can design inhibitor in biotech
* Enzyme inhibitors are designed bases on active site
* Atorvastatin – bound to coenzyme HMG – lowering cholesterol
* Development of drugs progress because of kinetic and structure study of the enzyme (the ones we studied last lecture, equations)
* Two classes of enzyme inhibitors
  + Reversible inhibitors – noncovalent bond
  + Irreversible inhibitors – covalent bond
* ES reaction
  + E + S forms ES and then form E + P
  + P doesn’t go back and form E + S
  + Delta G support direction of toward P
  + Inhibitors bind active site and prevent E from binding OR bind somewhere else and change the shape of the S
  + E + I then EI then no products
* Reversible inhibition
  + Many drugs follow this
  + 3 types – competitive, uncompetitive, non-competitive or mixed – different modes of activities
* Competitive inhibition – inhibitor blocks the pocket of the active site – substrate cannot bind so no get products – S can be processed but I cannot
  + In the absence of I – will follow hyperbolic curve in blue – reaction proceeds well even with little S
  + When add I – it competes with S to get to the active site – has good chance of getting to active site before S – slowing down the reaction – shape of curve doesn’t change but rising very slower – the curve moves down
  + 
  + In straight line plot, add I, rate of reaction goes down, slope increases – getting less product – KM increases a bit – vmax remains constant
  + Km increases so need more S to get to the same rate
  + Slope changes but intercept does not change – where it cuts the x-axis changes
  + 
  + Examples: (common questions) methanol poisoning, ethanol and methanol complete for the same action site – can cause blindness – when people drink unrefined wine – methanol forms formaldehyde - treatment is to give ethanol – when ethanol is bound, methanol is removed from the body harmlessly
  + FQA – use ethanol to treat methanol poisoning
  + Malonate competitively inhibits succinate dehydrogenase – S is succinate – P is fumarate – I is very similar to S – malonate binds to active site and stop succinate from binding
  + Drugs are designed to fit active site – eg. oseltamivir – Avian Flu – fit enzyme pocket very well – control spread of the virus
  + Cancers can be treated by I – if we prevent adenosine being processed to insosine
  + HIV enzyme inhibitors
  + ACE inhibitors – angiotensin I casues high blood pressure and vasoconstriction – cut by ACE at the proline – forming angiotensin II – prevent congestive heart failure – captopril binds strongly with the enzyme (nm) while angiotensin I is um
* Uncompetitive inhibition
  + Does not bind to active site – not compete with S – S binds
  + As S binds, I bind at the regulatory site –sit far from active site – shape of E modified – active site cannot make P
  + So bind after ES forms but prevent P formation
  + Called Dead-end complex – E is neutralised – just go to corner and wait until I go or stuck there forever
  + Normal way to prevent E from functioning – reduce S in the circulation
  + 
  + EIS – enzyme inhibitor substrate complex
  + Graph does not reach vmax – both vmax and kmax change
  + Vmax is affected more
  + Km shifts because vm shifts
  + With inhibitors, curve goes up – form parallel lines –simply move upward as more I
  + Slope remains the same but Km and Vmax decrease
  + 1 / [S] is increasing
  + 
  + Example: anticancer drug methotrexate – green is the enzyme – black is the drug molecule methotrexate – cofactor in purple – when I bind after S is bound – S is cofactor – no product forms – I is sitting where S would bound
* Mixed or noncompetitive inhibition
  + I bind either to E directly or ES complex
  + It does not compete for active site
  + No P forms
  + Effect of vmax is more than effect on Km
  + Km has only small changes – almost the same
  + Vmax is decreased
* 
* Irreversible inhibition – cause rapid paralysis
  + Eg. I prevent molecule from conducting nerve conduction – DFP binds to the active site serine of acetylcholinesterase – form covalent bond with Serine – releasing HF – HF itself is toxic
* Enzyme regulation is very important – how much E is present – speed or slow it down – can be done at 3 levels
  + Gene level control – multiple isomeric forms called isozymes
  + Faster – allosteric mechansims – structural changes
  + Covalent modification – serious changes caused by forming covalent bond – usually phosphorylation or dephosphorylation – on or off switches
* Isozymes - LDH is at the end of sugar – different forms are created and form tetramers (quaternary structure) – see different forms in different cell types – regulating by making different mixes of this E (isozyme)
* Production of different enzymes – have different E activity – different E population provides different regulation and rate – can lead to diff products being formed – pathway becomes deviated leding to diff sets of products
* Allosteric control – final product actually controls the activity of the E in the very first step
  + Eg. ATCase – when have sufficient amount of CTP – it inhibits ATCase in the first step – when it is low, reaction happens again
  + Quick and efficient way – change from hyperbolic to sigmoidal – can be in positive or negative direction
  + ATP is positive effector – CTP is negative
* Reversible Covalent modification – phosphorylation, acetylation, etc. – covalent modification to control protein activity – post-translation of aa
  + Phosphorylation is most important – carried out by protein kinases – attaching phosphate to OH – converated to PO3 – need ATP – P is also provided by ATP – switches on E
  + Switch off – phosphatase removes P and generate E with OH (original state)
  + ATP is phosphate donor
  + A lot of activating signals
* Enzyme mechanisms
  + Fischer in 1894 think of E as locks – unlocked by S – bound together because of shape complimentary
  + Allosteric modification – lead to induced fit model – when S binds to E – they both adjust to fit each other – slight medication leading to transition state – when P released E goes back to original state
  + Structure of proteins – real shift when S binds and no bind – eg. hexokinase when glucose binds the enzyme change shape to sort of hug the glucose