TOPIC 4: ANTIBIOTICS

October 26, 2020 4:40 PM

ANTIBIOTICS

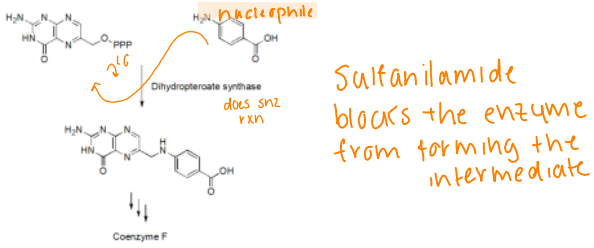
- a small molecule drug
- "anti-life"; selective poisons for microbes
  - bacteria (antibiotics), viruses (antivirals), fungi (antifungals), protozoa
- antibiotics have had largest impact on human life of all meds except vaccines
  - vaccines are the greatest achievement of modern medicine
- infectious diseases has been the leading cause of death throughout history
  - in middle ages 25% of europe killed by the plague; entire towns killed
  - everyday infections common cause of death
- life b4 penicillin in 30s/40s: could die from small cut infection so ppl were v scared
  - huge change in attitude today

ERLICH AND THE MAGIC BULLET

- paul erlich developed first concept of antibiotics in 1907
- called magic bullet bc drugs usually permeates through all cells in the body and spreads, but this was a drug that would get exposed to entire bo dy, but would only target the bacteria in body, not other cells
- trypan, a red dye effective against trypanosomes
  - the dye was attracted to the organism
  - trypanosomes are the infectious agent of african sleeping sickness
    - has made lots of africa unavailable for farming
  - red dye selectively colors trypanosomes (enter trypanosomes, but not RBCs)
- erlich thought this selective idea could be used to make a drug selective to trypanosomes that wouldn't enter other cells
- he used trypan as a lead and did SAR to optimize the material from dye→drug
  - arsenic chemically similar to nitrogen and its toxic
  - he though if he swapped the middle nitrogens of the dye molecule for arsenic, it would be selective for the organism and kill it
- became first ever antibiotic called salvarsan 606
- was actually a poor product; didn't sell well bc salvarsan wasn't druglike; not convenient
  - the drug worked well against syphilis (which was huge disease back then) but it was so inconvenient that it just didn't sell
  - weekly injections for months
    - drug v insoluble; need to use a v dilute solution
    - 600mL each injection
    - required 1 hour
    - syringe was used (slow i.v. drip invented 1935 after salvarsan)
    - intramuscular injection produced necrosis
      - v dangerous if injection wasn't done right
      - would kill the arm tissue/muscle and would need to amputate

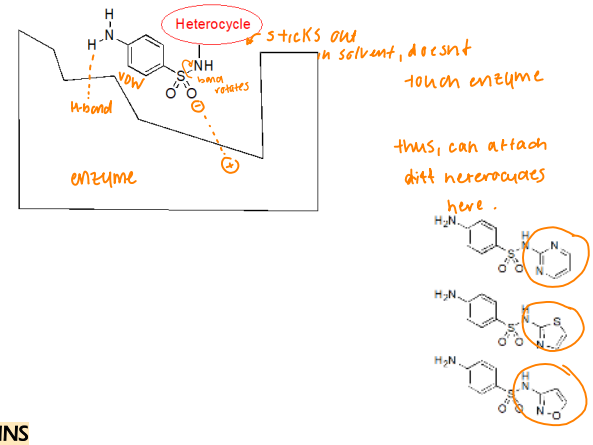
DOMAGK-SULFA DRUGS

- gerhard domagk at IG Farben 1932
- next antibiotic, which was a commercial success
- was investigating dyes to use as a lead and discovered prontosil, a red dye which could be effective in killing infections in mice
- domagk gave his daughter (who had a throat infection) this antibiotic and she recovered
- effective in vivo (in living systems) but ineffective in vitro (in glass)
  - prontosil requires metabolism to become effective
    - it gets cut in half, which makes the effective version of the drug (sulfanilamide)
- thus, prontosil is a pro-drug; its not a drug, but can become one
- they decided to sell sulfanilamide as the drug, not prontosil bc it was red and they found that animals that took it were tur ning red-bad side effect
- mechanism of action
  - bacteriostatic; interferes with growth
    - does not kill bacteria; only effective if patient has competent immune system
  - interferes with tetrahydrofolate (coenzyme F) biosynthesis; bacteria needs this to grow and divide



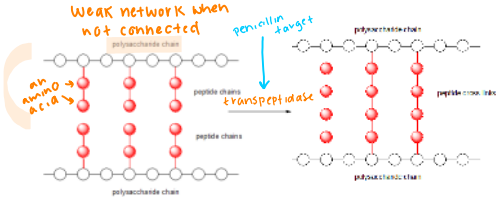
  - prevents bacteria from making coenzyme F; bacterial growth stops
    - sulfa drugs inhibits functioning of the enzyme dihydropteroate synthase which brings two components together via SN2 rxn to make the intermediate for coenzyme F
  - sulfonamides are competitive inhibitors of PABA (aka para amino benzoic acid)
    - sulfonamide competes with PABA for active site on enzyme
  - sulfanilamide has the same pharmacophore pattern as PABA
    - amino group on left (h-bonding), nonpolar region in middle, and, ionic bond on right (neg charge on molecule, pos charge in active site)
  - ideally the inhibitor binds to the enzyme better than the natural substrate does
    - this particular drug does actually bind better than the natural substrate
      - nature does not make optimal systems, just good enuf

        - if nature were optimal, drug wouldn't be possible (couldn't compete with binding pocket)
    - basis for selectivity in humans
        - humans lack the enzyme; the drug wont act on our cells, just the bacterial cells
        - get tetrahydrofolate from folic acid in our diet
- common structural modifications of the drug



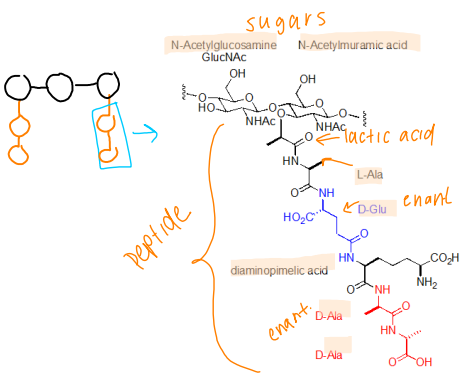
PENICILLINS

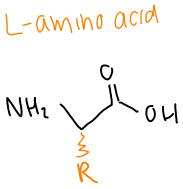
- 80% of all antibiotics. more than 30k synthesized, more than 100 sold as drugs
    - most are semi-synthetic; natural form of penicillin is not drug-like, need some synthetic sources to make modifications of the natural source penicillin
- alexander fleming contaminated his bacterial culture in 1929 with **penicillin mold** (this mold produces the drug)
    - there was zone of death of bacterial colonies around the mold on the agar plate
    - he reasoned that the mold was secreting a substance that was preventing growth
    - but what he actually observed on the plate was that out of 2 bacterial species, one did not grow near the mold while the other did
        - so he though the mold could be used as a purifying agent
        - missed the main point of using penicillin to treat infection
- florey and chain isolated penicillin in 1941
    - their goal was to use it as a drug
    - they infected mice with bacteria and gave half of them penicillin and half of them nothing
        - those given penicillin were alive, those that were not died
    - also did first human experiments using penicillin; in 1943, the drug was commercially available
- mechanism of action
    - penicillin has great selectivity for bacterial cells vs human cells
    - bactericidal; kills bacteria
    - bacteria have a cell wall; very rigid outer layer that gives strength
    - bacterial cells have internal pressure (one of the reasons they are so small compared to human cells)
        - the cell wall resists this osmotic pressure
        - there is a higher concentration of biological molecules inside the cell than outside where there is lower conc
        - this makes a **concentration gradient**; by le chatelier's principle, this wants to be equalled out
        - cant move stuff from in n out bc of the membrane, but can easily move water outside to inside to reduce the inside concentration
        - moving water into the cell increases the pressure inside the cell
        - large cells have a low internal pressure (human cells)
        - small cells (bacteria) have a high internal pressure; these cells req extra structure to contain the osmotic pressure (cell wall)
    - cell wall is made of peptidoglycan structure; polysaccharide chain with peptide cross links in between to attach the polysaccharide chains
        - these molecules are v large, so they are v hard to move-gives cell wall its strength (strength comes from the cross links)
        - formation of the cross links is the last step in cell wall biosynthesis
            - done by transpeptidase enzyme

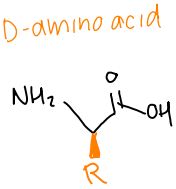


    - the polysaccharide chains are polymers: small molecules (monomers) linked to make giant polymers
        - properties are changed (compared to its monomers), polymers are very strong molecules because of their size (don't move around very much)
            - ex: resin monomers are liquid. once polymerized, becomes solid, rigid polymer (make plastic)
        - linear polymers give spaghetti-like mixtures; a gooey liquid
            - long molecules can slide past e/o; the structure these make is flexible bc the molecules can move relative to e/o
        - cross-linking is a method of strengthening structures
            - form connections bet strands (or even bet diff parts of the same strand)
            - this makes a rigid, strong 3D network in which the molecules cant move much
    - the cross link precursor has two types of sugars that alternate and link to make the polysaccharide chain
        - lactic acid (not an a.a., related to the sugar) is the connecting part bet the polysaccharide chain and the peptide
        - the peptide is composed of an a.a seq that has really unusual structural features (ancient structure of bacteria, not found in other living systems):
            - L-alanine a.a (normal)
            - D-glutamic acid; enantiomer of the glutamic acid normally found in proteins
            - diaminopimelic acid (special a.a)

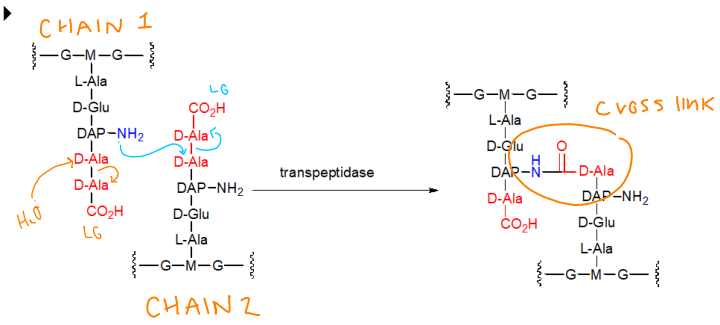
- two units of D-alanine (enantiomer of the normal alanine)
- another unusual feature is the crosslink bet diaminopimelic acid and D-Glu; the connection point is done using the side chain of D-Glu
- L-amino acid; enantiomer with side chain behind paper (dash)
  - almost all "natural" a.a acids have this configuration
- D-amino acid; side chain in front (wedge); rare in nature

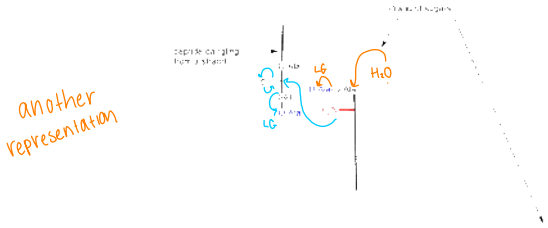


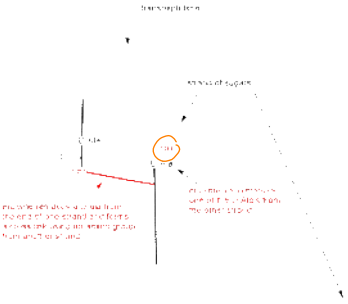




- transpeptidase does two related chemical reactions (nucleophilic displacements of D-Ala):
  - 1st rxn: takes amino group of diaminopimelic acid on chain 1 and uses it as a nucleophile to make a bond with the D-alanine of chain 2, LG leaves
  - 2nd rxn: takes a molec of H2O and the oxygen lone pairs act as a nucleophile and attack the D-alanine on chain 1 to give a LG that leaves

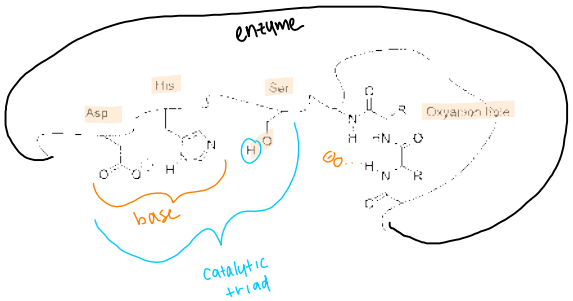


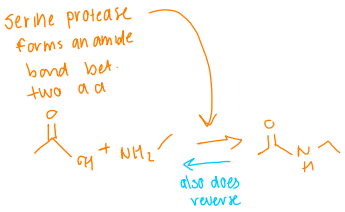




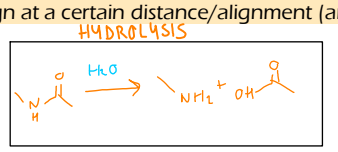
- transpeptidase is in a family of enzymes called serine proteases
  - refers to a structure and common mechanism by which the enzymes work
  - all enzymes in this family have a similar shape in their active site
  - serine proteases have four key features in their active site
    - a serine (OH) side chain
    - a histidine
    - aspartic acid in the negative form; acts as a base
      - histidine + aspartic acid together form a "super base"
      - purpose is to store the proton of the serine

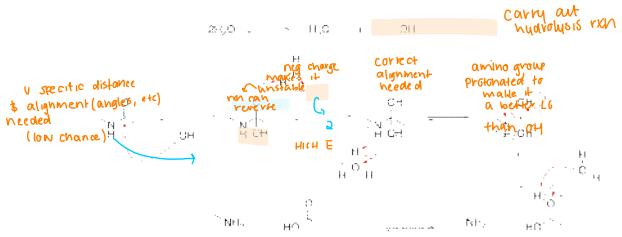
- "oxyanion hole" made up of amide bonds of backbone chain
  - generates an oxygen with a negative charge during the rxn
  - this negative charge is stabilized by h-bonding with NHs





- amide hydrolysis in solution is a very slow rxn; at pH 7.4, half life is hundreds of years
  - reason for slowness is bc for the rxn to occur, components must align at a certain distance/alignment (angle, trajectory)

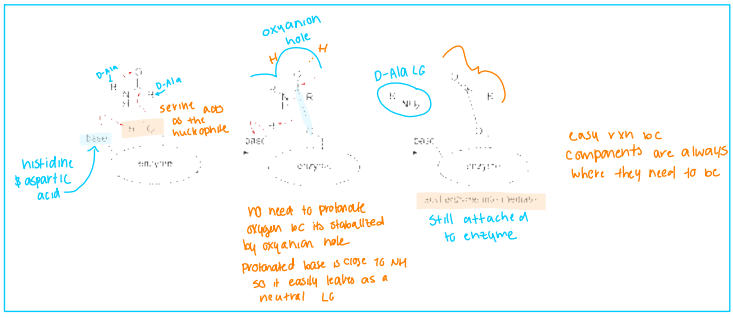




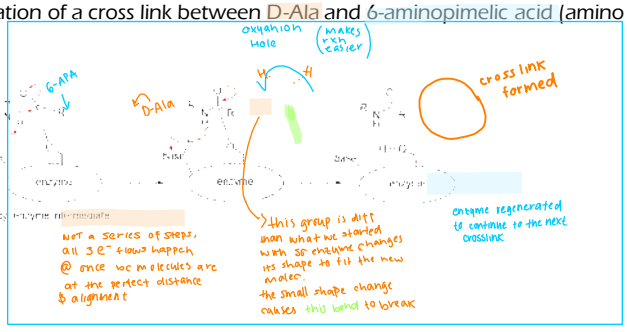
all steps in sequence are reversible
first 3 steps are v unlikely to happen

- amide hydrolysis in enzyme provides a perfect enviro for the rxn to take place
  - "designer solvent"; uses non-bonding interactions to make substrate go into active site of enzyme
  - solvents often participate in the rxn; by h-bonding, catalysts etc
    - enzymes do the same things, they facilitate the rxn via h-bonding for example
    - enzymes do this more efficiently than solvents bc they have components lined up perfectly with substrate while solvents rely on random motion, rxn can happen in a second

FIRST RXN BY TRANSPEPTIDASE IN CROSS-LINK FORMATION

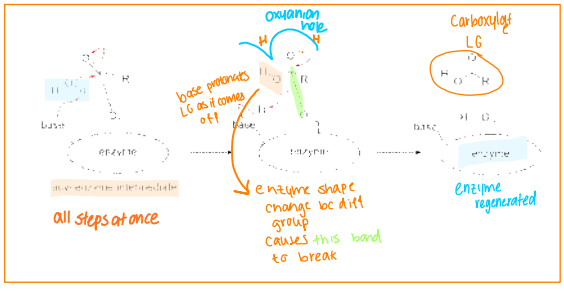


  - the enzyme will then "turn itself over" to detach from the molecule; transpeptidase can use two diff nucleophiles to break apart the acyl-enzyme intermediate
    - first is formation of a cross link between D-Ala and 6-aminopimelic acid (amino group acts as nucleophile)

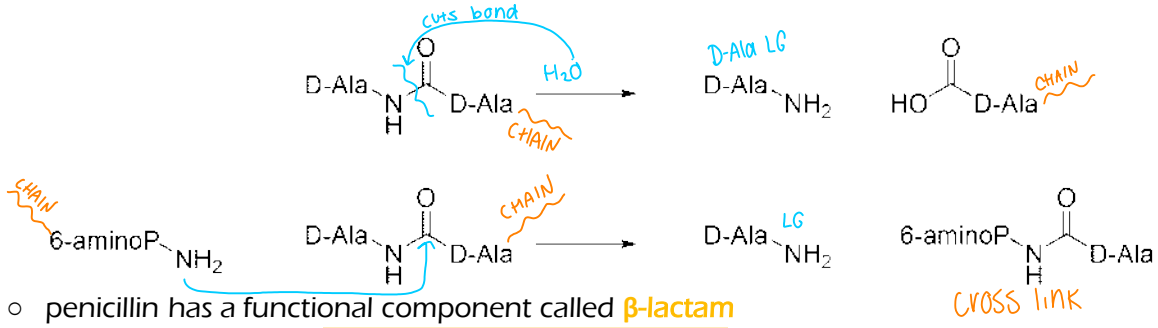


    - the second method is with water as the nucleophile; hydrolysis of one D-Ala-D-Ala peptide bond

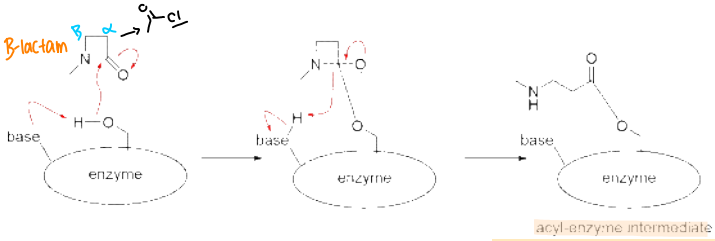
SECOND RXN BY TRANSPEPTIDASE IN CROSS LINK FORMATION



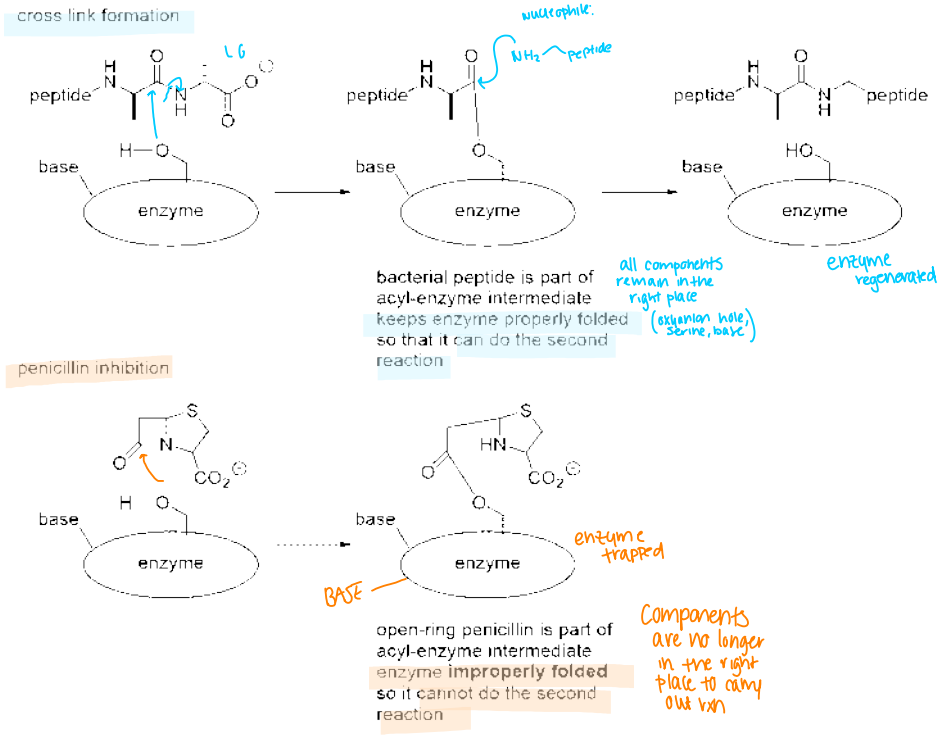
- **SUMMARY**: reactions catalyzed by transpeptidase
  - cross link formation
  - "trim" one D-Ala from each strand



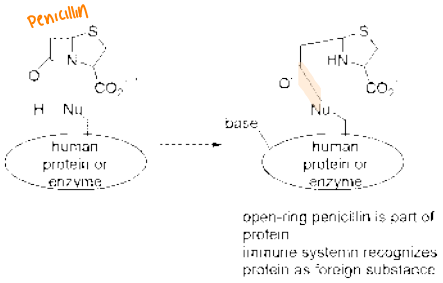
- penicillin has a functional component called **β-lactam**
  - a β-lactam is an amide in a four-membered ring
  - nitrogen is connected at β carbon
  - the ring acts as an electrophile (similar to acid chloride)
  - the transpeptidase enzyme opens the β-lactam ring



- when penicillin reacts with transpeptidase, the acyl-enzyme intermediate formed is the wrong shape
  - the enzyme gets stuck at this stage, can no longer be used to form crosslinks
  - irreversible inhibition (inhibition is permanent)
    - enzyme is unable to hydrolyze the ester formed
  - cell wall formation is prevented
    - cells burst from turgor (internal pressure) when dividing
- penicillin finds the active site through non-covalent interactions
  - right shape to fit in the active site pocket

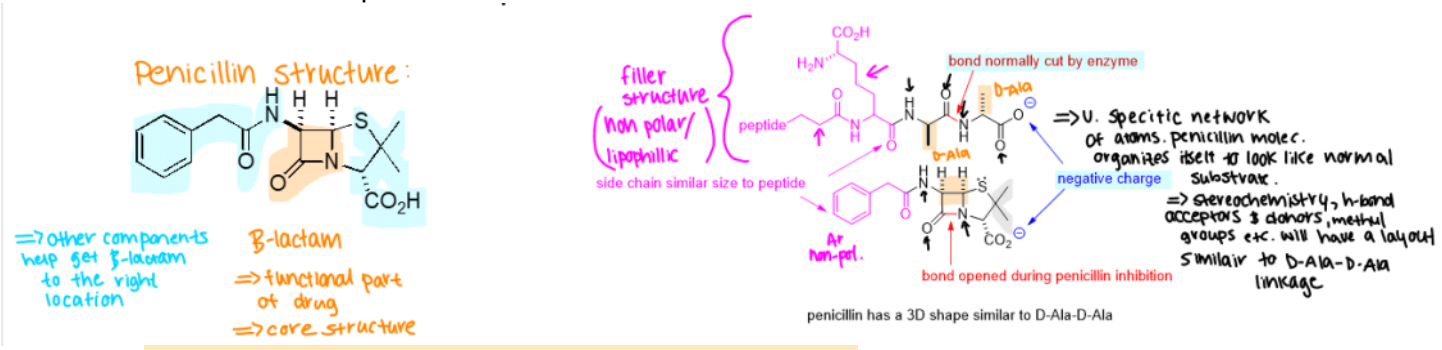


- selectivity
  - humans don't have cell walls; no equivalent enzyme
  - D-Ala-D-Ala linkage is unique in nature
    - bacteria only, virtually all other a.a. have L config
  - very clean drug
    - non-toxic, safe
    - few side effects; major side effect is allergies (~0.1% incidence)
  - source of allergy
    - penicillin is a good electrophile
    - can acylate the host serine proteases or other proteins
      - serine proteases w/ larger than normal active site that can fit penicillin
      - proteins w/ unusually nucleophilic side chain
    - acylated enzymes/proteins are recognized as foreign by the immune system and get attacked
    - strongest rxns happen second time taking the drug

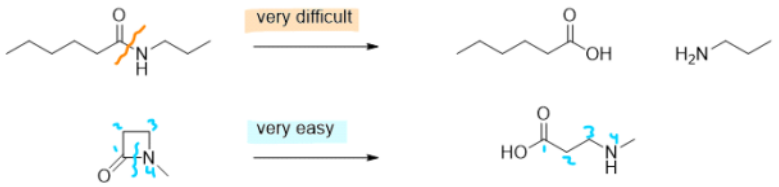


PENICILLIN DESIGN

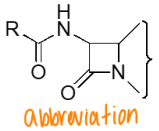
- penicillin gets into the transpeptidase active site by using a network of non-bonding interactions that have the same 3D shape as the network of the normal substrate

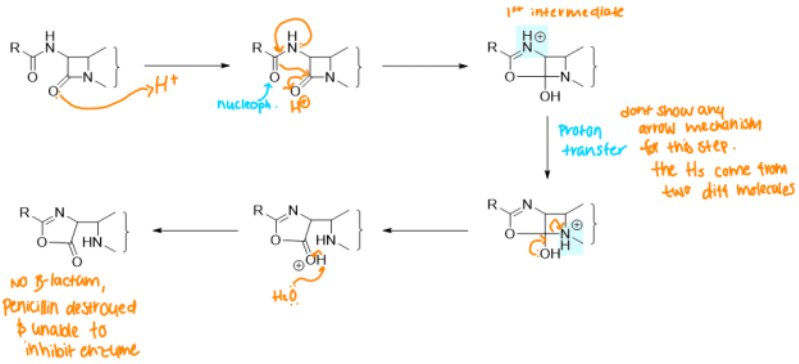


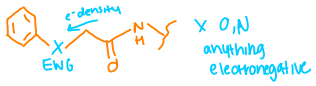
- thus, penicillin is attracted to the active site of transpeptidase and β-lactam is placed in the correct position for the enzyme to cut the bond
- amide bonds are usually very strong, but β-lactam bond is actually very easy to break due to its ring strain:

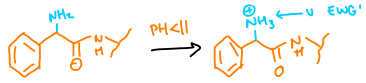


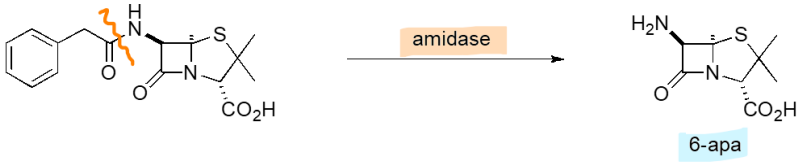
- penicillin G is the natural form of penicillin isolated from mold
  - it has 4 significant limitations as a drug
    - acid sensitivity; if pH is too low, the molecule falls apart
    - resistance; certain b
    - bacteria cant be killed by penicillin G
    - spectrum of action; there are two major families of bacteria, penicillin doesn't work well on one of these
    - bioavailability

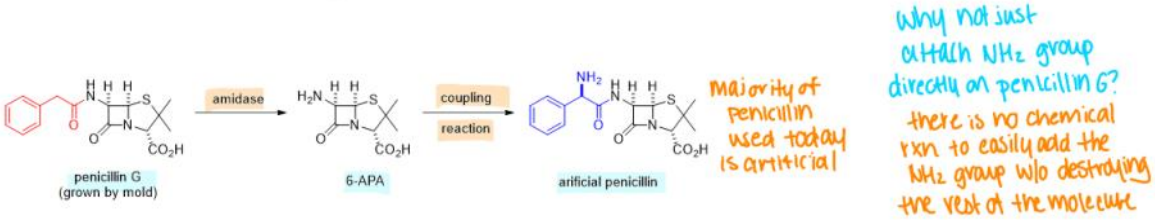


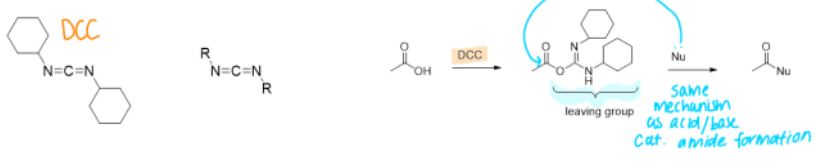
- acid sensitivity; stomach environment is very acidic so drug must be given by iv injection
  - impossible to make pill form of penicillin G
  - reason for acid sensitivity:
- 

- engineering for acid resistance: if we make the nucleophilic oxygen less nucleophilic, further steps in the rxn can be prevented
  - add electron withdrawing group next to the oxygen
    - can add a heteroatom between the oxygen and benzene
    

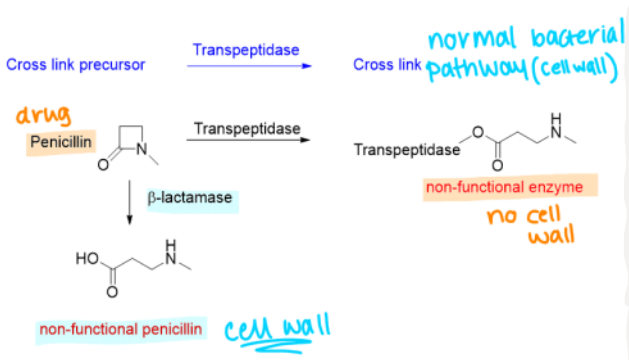
    - can add a positive charge next to the oxygen
      - an NH_2 group at pH<11 (body pH) will take on a positive charge
      

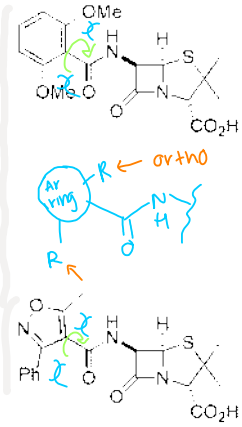
- the penicillin core is v complex and hard to synthesize, so the drug is made via **semi-synthesis**
  - the core (6-APA) is obtained from biological source
    - extract penicillin G from penicillin mold
    - use amidase enzyme from e.coli to remove side chain
      - modern versions have been optimized by genetic modification; use modified bacteria that have been engineered to be very efficient at doing this
      

  - the new drug is made with synthetic transformation
    - convert penicillin G into 6-APA like above ^^
    - attach acyl group to NH_2 of 6-APA
    

      - the coupling reaction takes a carboxylic acid and connects it to a nucleophile
      - DCC is a peptide coupling reagent
        - DCC is now more commonly used in proteins, but was originally developed for coupling of groups onto penicillin
      - the acid is mixed with DCC (don't need to know mechanism), this makes a leaving group, nucleophile attacks
      

- resistance; many bacteria develop resistance to penicillin
  - one reason is mutations in transpeptidase that result in penicillin not fitting in the active site (enzyme structure changed)
  - another way is the bacteria evolving ways to get rid of the drug
    - ◇ actively transport penicillin out of the cell
    - ◇ destroy the drug
  - bacteria started to produce the enzyme β-lactamase aka penicillinase (a serine protease)
    - ◇ this enzyme evolved to be optimized to open the β-lactam ring of penicillin
    - ◇ bacteria use transpeptidase to make cell walls and β-lactamase to protect against penicillin
    - ◇ bacteria secrete β-lactamase just outside the cell, as penicillin approaches the cell, it the enzyme opens the ring, making the drug non-functional
    - ◇ first appeared in *Staph. aureus* in the 60s
  - when resistance arises, how does the drug industry overcome it?
    - ◇ designed a diff version of the penicillin; the penicillin regains activity using a "shield"
    - ◇ if you put a large substituted heterocycle on the left, it allows you to kill bacteria that are normally resistant to penicillin
      - □ the heterocycles have substituents at both ortho positions
      - □ the large heterocycles will adopt a conformation that wont allow it to enter the active site of β-lactamase; the shield provides selectivity





Methicillin

Oxacillin

=> will adopt a 3D conformation that makes the structure v big & unable to fit β-lactamase pocket:

double bond character due to resonance

=> these bonds will be flat (on the same plane)

=> so this entire structure will be lying on the plane of the paper

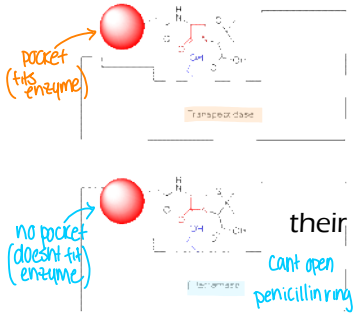
=> this flat structure is okay if there are just hydrogens @ ortho posn but other groups will have steric interference

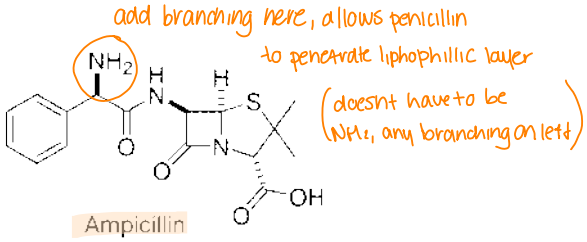
=> bond will rotate 90° bc sterics

=> the ring is not flat anymore, v big 3D struct

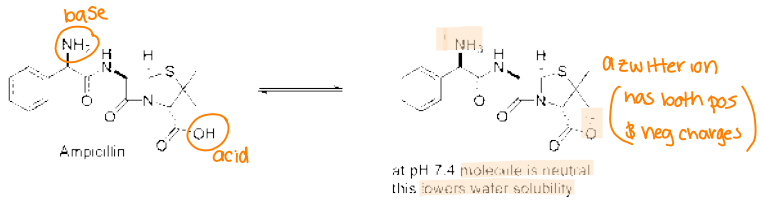
=> transpeptidase has a big active site pocket, so this big structure will still fit

- spectrum of activity
  - two general types of bacteria:
    - ◇ gram positive (blue)
    - ◇ gram negative (red)
  - original penicillins were only active against gram positive strains
    - ◇ narrow spectrum of activity
  - desirable to have activity against gram negative strains; want a wide spectrum of activity
  - gram positive bacteria have their cell wall on the outside
    - ◇ easy for penicillin to get to transpeptidase at the cell wall and stop the rxn
  - gram negative bacteria have a lipopolysaccharide layer on the outside, then cell wall
    - ◇ bacteria are coated in a specialized layer of "slime"; extremely lipophilic
    - ◇ penicillin cant penetrate this layer and get to the cell wall
  - after experimentation, found out that putting branching on the left side of penicillin molecule allows it to penetrate the outer lipopolysaccharide layer and inhibit transpeptidase

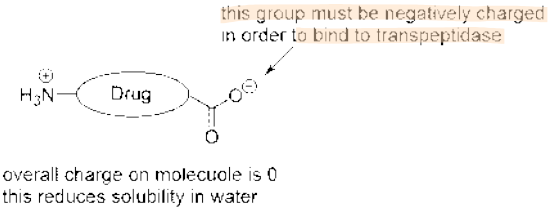


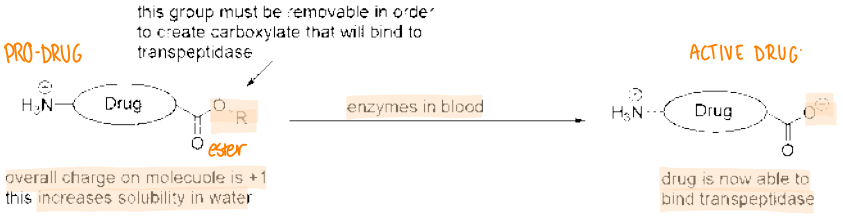


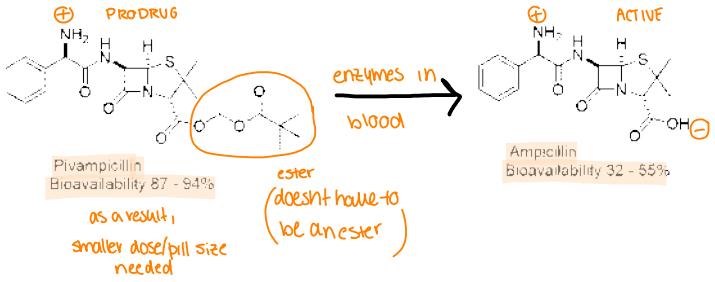
- bioavailability; some forms not very water soluble
  - results in poor bioavailability (poor absorption from intestine)
    - ◇ must be in solution as it approaches the membrane of intestinal cell
- need to improve water solubility; need an overall charge
  - penicillin as it is has a net neutral charge



- the molecule must preserve the negative charge on the right in order to bind to transpeptidase
  - need negative charge on right as penicillin molecule approaches bacteria
- prodrugs; add removable groups to a drug to improve absorption in the body
  - once in the blood, this group is removed to deliver the drug in active form
  - a prodrug is used to get penicillin into body, then gets changed by metabolism into the active form of the drug

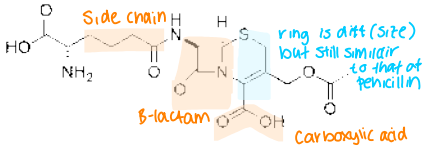




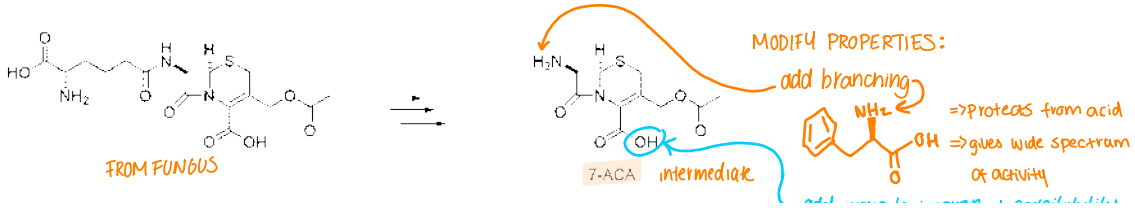


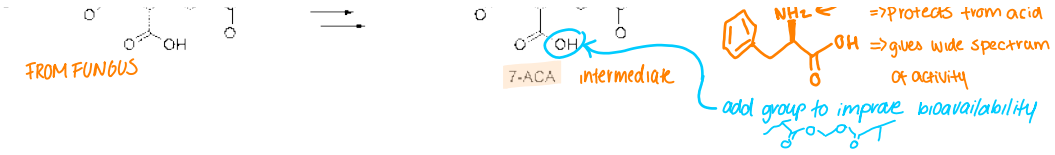
RELATED ANTIBIOTICS

- after penicillin, scientists tried to find other antibiotics
  - cephalosporins were discovered in 1948 in fungus growing in an italian sewer
    - ◇ structurally similar to penicillin

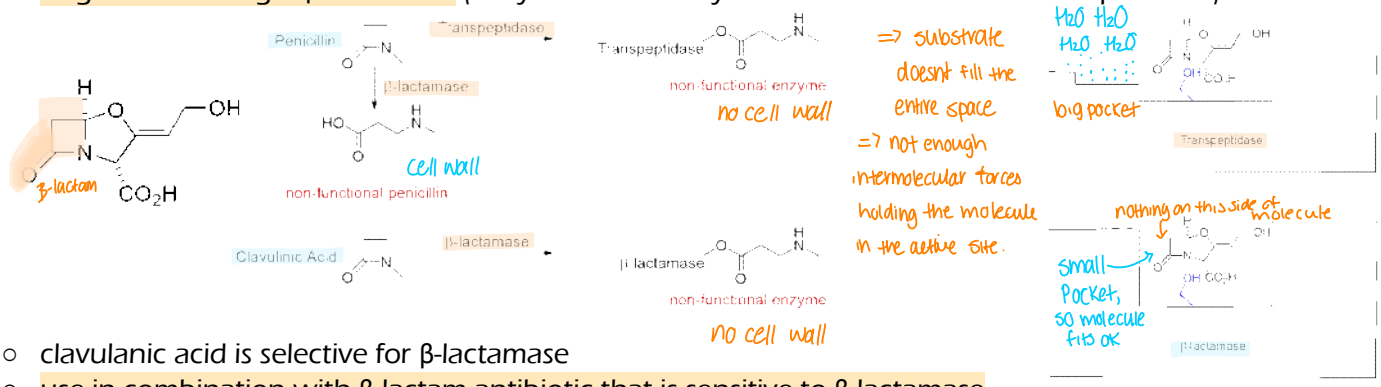


    - ◇ kills bacteria using same mechanism as penicillin
      - □ blocks action of transpeptidase by forming a covalent bond with the enzyme
    - ◇ cephalosporins have some advantages over penicillin
      - □ its more lipophilic, so has a broad spectrum of activity w/o modification
        - ◆ kills both gram neg and gram pos bacteria
        - ◆ gave scientists a clue of how to modify penicillin to kill gram neg bacteria
      - □ [4,6] (cephalosporins) ring systems are less reactive than [4,5] (penicillin)
        - ◆ shows activity against some penicillin resistant bacterial strains
          - ◇ harder for β-lactamase to open up the ring
          - ◇ less risk of allergy; less chance it reacts with a nucleophile
    - ◇ disadvantages compared to penicillins:
      - □ not orally active bc sensitive to acid (same reason as penicillin acid sensitivity)
      - □ bc its less reactive, its less potent so need larger doses
    - ◇ used semi-synthesis to modify structure of drug to tune its properties:

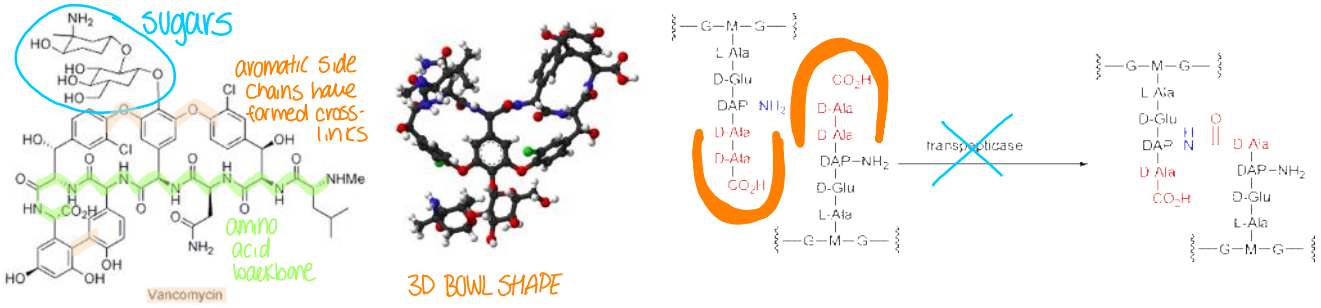



FROM FUNGUS
7-ACA
intermediate
NH2
OH
=>protects from acid
=>gives wide spectrum of activity
add group to improve bioavailability

- clavulanic acid discovered in 1976 in a species of bacteria
  - NOT an antibiotic
    - ◇ does not inhibit transpeptidase; does not kill bacteria
  - target of this drug is β-lactamase (enzyme secreted by bacteria that are resistant to penicillin)



  - clavulanic acid is selective for β-lactamase
  - use in combination with β-lactam antibiotic that is sensitive to β-lactamase
    - ◇ clavulanic acid inhibits β-lactamase and protects the antibiotic from the resistance enzyme
    - ◇ the antibiotic can then inhibit transpeptidase and kill bacteria
- clavulanic acid is an example of a drug-drug interaction
  - one drug changing the bioavailability of another is the most common drug-drug interaction
  - ex: grapefruit contains a material called bergamotin which deactivates key enzymes in the liver
    - ◇ the molecule is metabolized by the liver and the metabolite irreversibly binds a glutamine on CYP_{450} and deactivates CYP_{450}
    - ◇ if you are taking a drug and eat grapefruit, could have a toxicity issue
      - □ normally, when making a drug, the pill is made with a bigger dose knowing some of it will be metabolized by the liver and only a portion of the drug will be used by the body
    - ◇ if eat grapefruit, none or less of the drug is metabolized and you are taking a bigger dose of the drug than intended, can lead to overdoses
- vancomycin was discovered in 1956 from bacteria
  - a glycopeptide drug; formed from a short protein (peptide) sequence containing 7 amino acids
  - the amino acids are chemically modified with sugars (glyco)
  - the amino acids are aromatic; these rings are cross linked and form a bowl-shaped structure
  - blocks the formation of bacterial cell walls
    - ◇ vancomycin binds strongly to the D-Ala-D-Ala tail of peptide chains
    - ◇ prevents transpeptidase from forming cross links to make the cell wall
  - enzyme inhibition by substrate binding



  - not a widely used drug, its a last resort
    - ◇ used for infections with bacteria that are resistant to other antibiotics
    - ◇ limitation: works only on gram positive bacteria bc vancomycin is too large and not lipophilic enough to cross the lipopolysaccharide of gram neg bacteria
- most families of antibiotics were discovered in 40s and 50s, the last one discovered was in the 90s
- companies don't do a lot of research on antibiotics
  - one of the main reasons is $$, companies stopped looking for natural products bc they dont make a lot of money from it
- modern antibiotic research is v limited; academic and occasionally specialized small company only
- antibiotics are too cheap and readily available
  - new drugs wont make them money

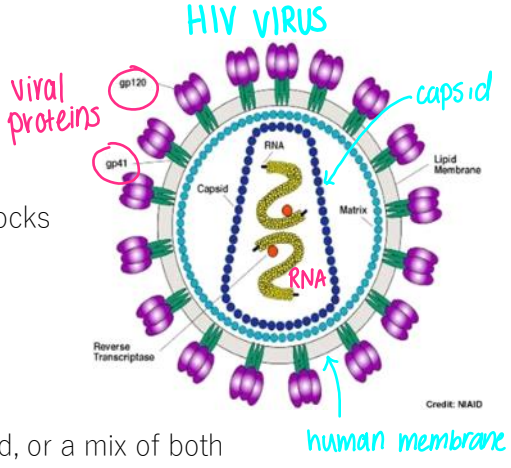
- new drugs are not likely to be used by doctors

TOPIC 5: ANTIVIRALS

October 26, 2020 8:24 PM

INTRO TO VIRUSES

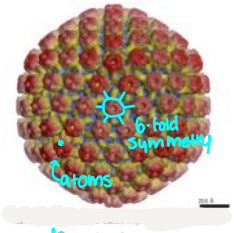
- provide a window into the functioning of biochemical processes
  - bc they are so simple (most viruses have only 10-12 genes), it reduces the complexity of studying how they work
  - still have complex behaviour with such few genes (ex: do multiple jobs with the same protein)
- impo for therapies; ex: cancer treatments and understanding how it works come from studying viruses
- most infectious diseases in NA is caused by viruses; more than 95% of all respiratory disease
- all class 4 pathogens are viruses; thus v deadly
  - ebola, marburg, lassa fever, hantavirus, smallpox
- classification of pathogens (dictates protective measures needed in lab)
  - CLASS 1: no risk or limited risk
    - work on open lab bench
    - P1 lab
    - ex: e. coli
  - CLASS 2: moderate risk
    - limited access to lab (password needed)
    - lab coat req
    - laminar hoods used; air only goes in, not out
    - P2 lab
    - ex: herpes
  - CLASS 3: risk of death or serious illness
    - restricted access, special training required (gov cert)
    - surgical gowns, gloves (sealed), respirator
    - all liquids/air coming in/out is filtered/treated
    - everything coming out (devices) is autoclaved and incinerated
    - ex: HIV, y. pestis (plague)
    - P3 lab (usually located in a P2 lab)
  - CLASS 4: lethal, highly infectious, untreatable
    - lab accessed by airlock, special training & cert req
    - space suit worn, shower w/ suit going in and out
    - low pressure in lab (therefore the leaks will be out to IN), airlocks
    - all liquids and gases filtered/treated going in/out
    - anything in/out is autoclaved/incinerated
    - ex: ebola, marburg, smallpox
    - P4 lab



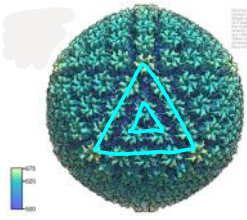
- virus structure
  - genetic info (DNA or RNA); can be single-stranded, double-stranded, or a mix of both
  - surrounded by a capsid; hollow container made of protein
  - some viruses carry additional proteins
    - enzymes; can help viruses get into the cell, take over the cell, and destroy its defence systems
    - regulatory proteins; have the ability to bind to proteins in the host cell and change their function. make sure viruses reprod, not the host cell
  - some viruses are enveloped
    - capsid is surrounded by a membrane; a remnant of the host cell membrane
    - viral proteins are embedded in the membrane; these have pieces on the inside that recognize the capsid

- prof thinks viruses are self-replicating piece of cellular machinery (not alive)
- the virus seems to be co-opting pieces of existing cellular machinery for the proteins
  - many viral proteins are structurally related to the host proteins
  - virus structure is v simple, it doesn't have the ability to exist by itself
  - can find certain oncogenes in viruses (proto-oncogene→oncogene (cancer) in humans)
- viral proteins often bind to host protein and alter that protein's function
  - ex: if one of our enzymes usually does something on our human proteins and we get infected by virus, a viral protein can stick to the host enzyme and cause it to carry out a rxn on a viral protein instead
- viral proteins tend to be organized around human cellular regulatory systems
  - viral proteins tend to resemble human proteins involved in nucleic acid replication and protein manufacture
- two common elements (all viruses)
  - duplication of genetic info of virus
  - production of viral protein
- both of the viral processes ^^ utilize host proteins (enzymes) and machinery
  - ribosome & nucleic acid polymerases
  - biggest challenge when developing drugs to treat viral infection
  - ribosome and polymerases are v picky in terms of what they will accept
  - bc viruses use these two to do things for them, its hard to target viruses bc it will interfere with normal human system
  - thus, lots of viral drugs have bad side effects (not v clean drugs)
- many viruses can be crystallized, can get an XRAY crystal-like structure to det 3D structure of the virus on an atomic scale

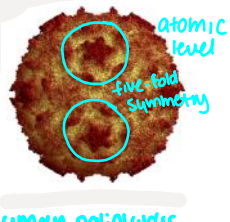
See v nice symmetry of capsids



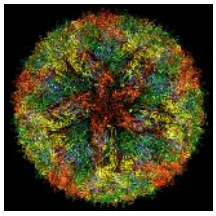
herpes virus
=> diff colors = diff protein



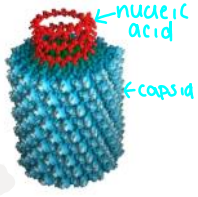
Cytomegalovirus
=> icosahedral virus
=> 20 triangular sections



human poliovirus



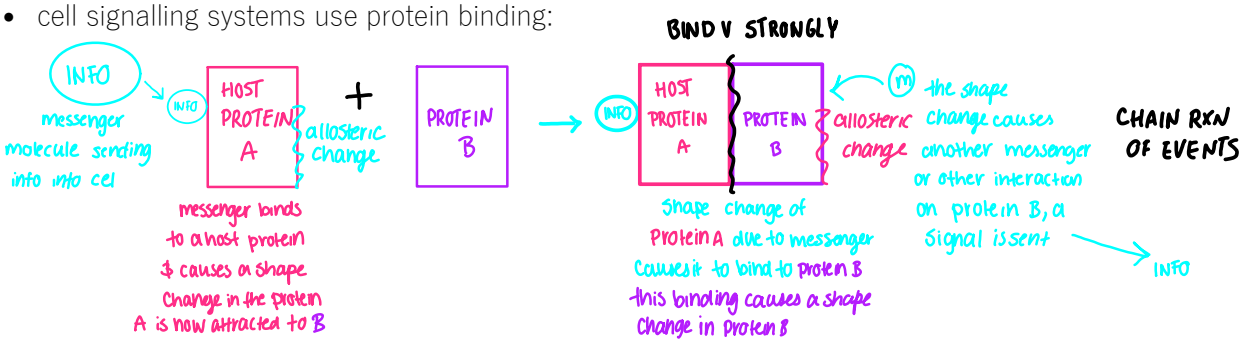
Zika virus
=> ribbon diagram



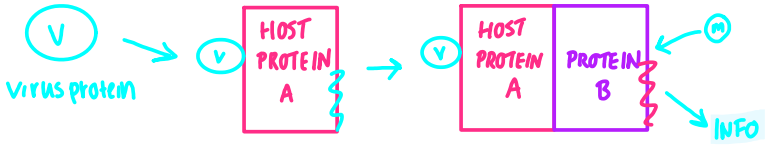
tobacco mosaic virus
=> can make this in lab by mixing nucleic acid & viral protein

CHALLENGES TO ANTIVIRAL DISCOVERY

- each virus is unique so each virus req a different drug
  - contrary to bacteria where penicillin can kill many diff bacteria person is infected by
  - must know what virus is infecting the person
- most viral proteins act by binding to host proteins
  - viral proteins are not enzymes, they are regulatory proteins
  - this is v hard to target w drugs
  - proteins are v big molecules and tend to bind to e/o v. tightly
  - need a large drug bc proteins are v. big and bind together over a large surface area
    - ◇ remember (by lipinski's rule): large drugs are not v. soluble, have a hard time passing through membranes; generally they are bad drugs
  - need to avoid interfering w/ normal host cells
- cell signalling systems use protein binding:



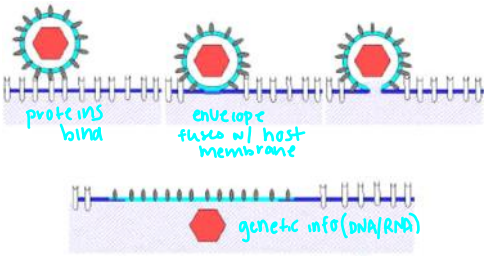
- many viral proteins also operate by binding to host protein
  - activates protein A when normal info isn't there:



  - controls the operation of the cell this way; such viral proteins are called **regulatory** bc they regulate the function of other proteins in the system
- drug has to get to the interface bet viral protein and host protein A and make sure it cannot bind to the host protein
  - drugs are normally designed to bind to the interface of the viral protein that normally binds to protein A
    - surface is no longer available to bind to the host protein
  - problem is that the surface is large, the drug must be a large molecule
  - large drugs work in the lab, but is hard to get it to work in humans bc bioavailability
    - insoluble and impermeable
    - MW>500 and likely too many HBD and HBA (violates lipinski's rules)
  - these types of drugs are limited for very serious/special cases and usually administered by injection
- viral enzymes as drug targets
  - easier than targeting viral proteins
    - most enzymes have small substrate molecules
    - can make small drug
  - but viruses carry few enzymes (usually 1 or 2-out of their 10-12 genes)
    - limited number of targets
  - most involve nucleic acid replication
    - host cells also make nucleic acids
    - substrates, mechanisms and structures are similar
    - the active sites are thus similar
  - so a drug that blocks a viral enzyme likely will also block a human enzyme
    - interfering w our enzymes that deal w nucleic acids is bad (side effects)
  - today, only a few good antiviral drugs exist
    - hepatitis C (cure)
    - herpes (ex: cold sores) (treat; life long disease)
    - HIV (mange; life long disease)
- major problems with antiviral drugs
  - selectivity; kill virus w/o killing host
  - diagnosis
    - drugs specific for each virus
    - diagnosis is an issue bc many viruses produce similar symptoms
      - more than 200 viruses cause colds
    - only way to know for sure is biochemical test
      - this takes time and drugs are most effective when used within first few days of infection
  - resistance
    - mutation rates in viruses very high
    - viruses quickly develop resistance to drugs (days or weeks)
    - ex: a drug for covid may not work anymore in 6 months
- immunization has been very successful for viruses
  - all time greatest achievement in medicine
    - smallpox (disease eliminated bc of immunization)
    - polio (disease almost eliminated)
      - 33 cases in the world as of 2018 (bc antivaxxers)
    - measles (rare)
    - mumps

THE VIRUS LIFE CYCLE

- virus life cycle; all viruses go through this sequence (some mix and match/blend these steps):
  - adsorption (sticks to cell via intermolecular forces) and penetration into cell (in a variety of ways)
  - capsid opens releasing contents (genetic material)
  - synthesis of regulatory proteins aka early proteins
  - synthesis of RNA or DNA
  - synthesis of structural proteins aka late proteins
  - assembly of viral particles (viral pieces are attracted to e/o)
  - release from host cell
- adsorption and penetration
  - adsorption: virus binds to host proteins on outside of host membrane
    - mediated by intermolecular forces bet proteins on virus and host proteins on membrane
    - ex: capsid binds directly and is passed inside before opening
    - ex: an envelope fuses with host cell membrane releasing capsid inside the cell
  - genetic info is injected into the cell
    - virus may also inject viral proteins (enzyme or regulatory proteins)
    - serves as a template for manufacture of early proteins



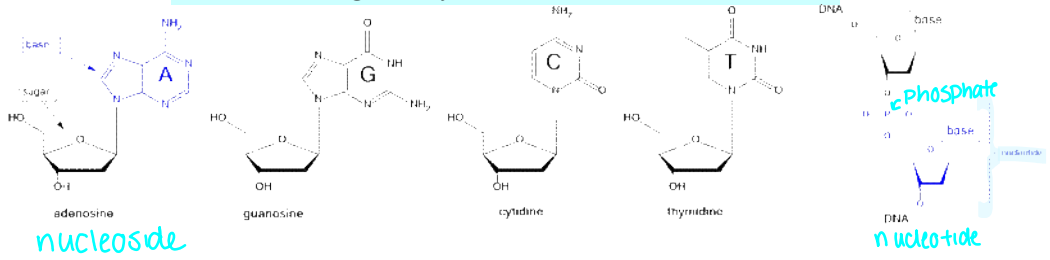
  - this is a poor drug target area
    - binding involves protein-protein interactions
      - protein surfaces are v large; hard to inhibit protein-protein binding with small molecules
    - limited success in HIV with a small molecule
      - fuzeon; peptide drug w 36 amino acids (large molecule)
      - maraviroc; small molecule drug
        - HIV has a protein on the outside that functions like a drill; the protein untwists itself when it binds to membrane and drills itself through cell
        - this drug can bind to this viral protein and makes it unable to unwind
- release of viral nucleic acid
  - capsid opens releasing genetic info into the cell
  - hard to target (involves protein interactions and pH changes)
  - two successful drugs; lucky accidents
    - influenza; amantadine and rimantidine (have v special structures, like detergents)
    - block an ion channel on the capsid of influenza
- synthesis of regulatory proteins
  - early viral proteins are made by the host cell using viral genetic info as a template
    - certain viral nucleic acids have sequences that allow them to directly enter ribosomes and cause expression of early proteins
  - build up of these regulatory proteins initiates take over of normal cell systems
    - viral nucleic acid replication
    - expression of viral protein
    - suppression of host cell defenses by viral proteins
      - certain viruses can prevent apoptosis
    - all involve binding to host proteins
  - no drugs currently exist for this phase

- synthesis of RNA or DNA
  - viral genome is replicated using host enzymes (polymerases)
    - RNA or DNA depending on the virus
  - some viruses have their own enzymes for this (drug targets)
  - most anti-viral drugs target this phase
    - require unique viral enzyme
      - if the virus uses only host polymerases, cant make a drug
    - prevent viral nucleic acid synthesis
- synthesis of structural proteins
  - utilize the host ribosome; poor drug target
  - some viruses utilize specific enzymes for protein maturation
    - the ribosome just makes one long protein chain of all the subunits
    - virus uses a protease to cut this into parts and release the individual proteins
    - protease drugs; HIV, hepatitis C
  - COVID also uses a protease
- assembly and release
  - capsid proteins self-assemble
    - nucleic acid inside
    - viral proteins outside
  - release may destroy the host cell
    - lytic virus
    - ex: herpes (skin cells), influenza
  - cell may remain intact; cell becomes a living factory for the virus
    - papilloma, herpes (in nerve cells; carry it your whole life)
  - only a few drug targets (HIV, influenza (not a v good drug tho)
- antiviral drugs require viral enzyme targets
- enzyme should structurally unrelated to host enzyme; provides selectivity
  - need diff binding pockets
- most viral enzymes are involved in nucleic acid replication
  - problem: viral nucleic acids are the same as human nucleic acids

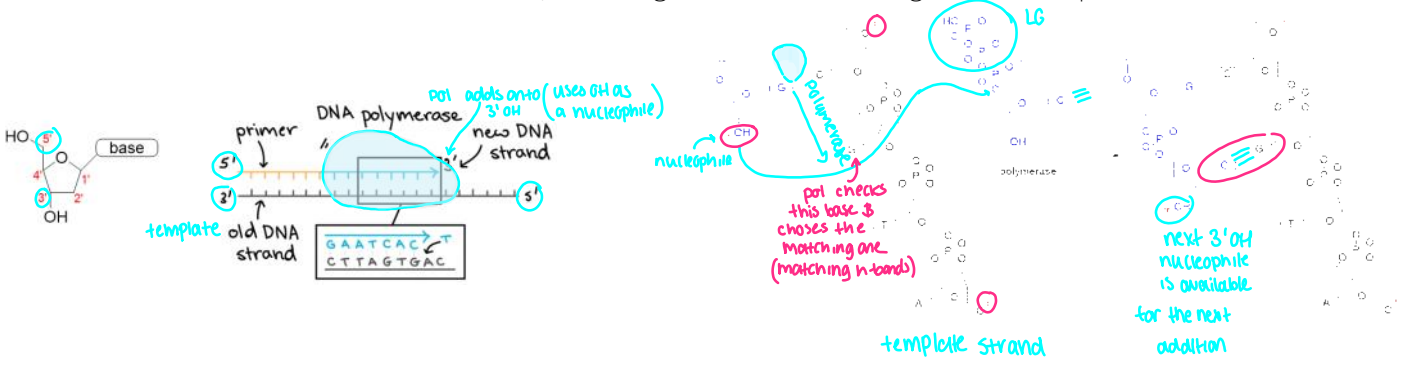
HERPES

- a family of viruses that causes chronic recurrent infection
- virus is able to escape immune system; latency in neurons
  - once in awhile there will be a stimulus, and the virus reactivates
  - exits neuron and causes an outbreak
- HSV-1 (primarily mouth/nose infections) & HSV-2 (primarily genital infections); the other subtypes are less prevalent
- HSV-1
  - cold sores and fever blisters on the mouth and nose, sometimes eye (serious)
  - more than 80% of the population is infected
  - only 10%-20% experience outbreaks
  - virus escapes eradication by latency inside neurons
    - inaccessible to immune system when latent in neurons
    - stimuluses like stress, sunlight, & immune suppression trigger outbreaks
    - virus travels down axon to epithelial cells the neurons are connected to and causes outbreak
  - lytic infection; only in epithelial cells, does not damage neurons
    - blisters last about a week but viral activity is short (less than 24hrs)
    - viral damage is
    - most damage is caused by the immune system
      - over stimulated and destroys most of the tissue (causes blisters)
      - drug must be administered quickly (first couple hours of outbreak to stop virus)

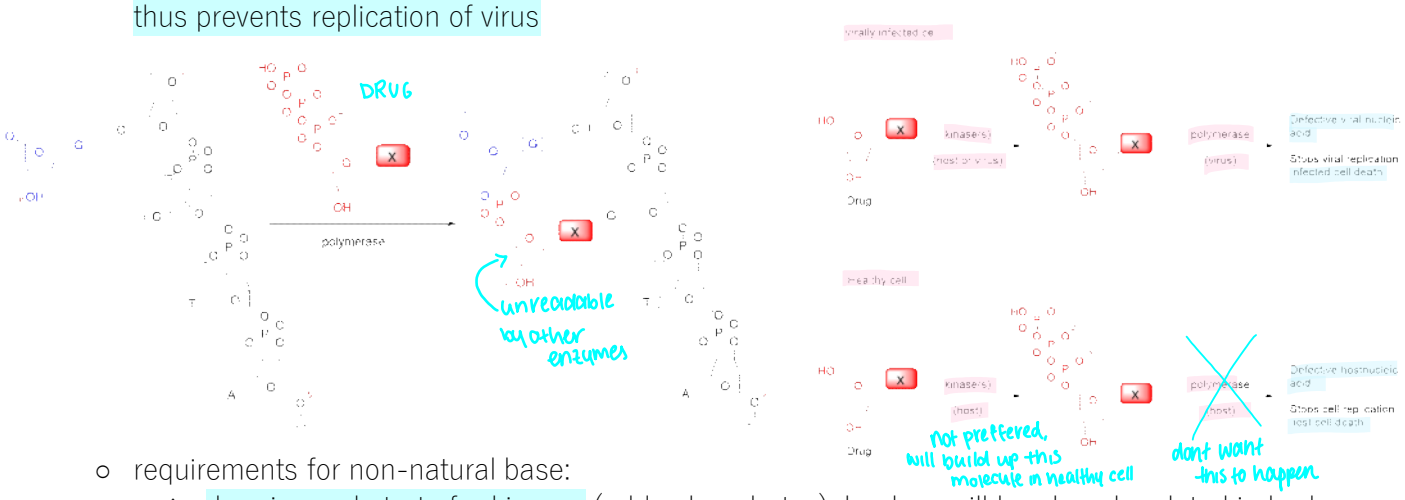
- HSV-2
  - sores and fever blisters on the anus and genitals
  - infects 15%-20% of the populations
  - infection is much more virulent and painful than HSV-1
  - usually one outbreak per year
  - most common type of STD
  - viral activity is short; viral damage is v minimal
  - most damage by immune system
    - drug must be administered v quickly to prevent outbreak
- recently, we see transmission of HSV-1 in lower body and HSV-2 in upper body
- herpes virus structure
  - genetic info is double-stranded DNA
  - v. complex virus
    - more than 70 genes (most viruses have <10)
    - virus has its own polymerase (drug target)
- polymerases assemble components of nucleic acids to make DNA/RNA:
  - nucleosides (no phosphate): made up of base and sugar
    - nucleotides have a phosphate on the sugar's OH
  - sugar (2-deoxyribose in DNA or ribose in RNA)
  - base: nitrogen containing aromatic heterocycles
    - the varying component of DNA/RNA
    - four diff bases (ATCG; U instead of T in RNA)
    - each base is recognized by other molecules bc of its shape and h-bonding pattern
    - enables binding of only certain structures to each base



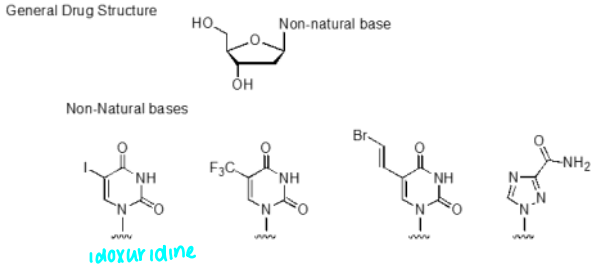
  - polymers are formed by connecting nucleosides together using phosphate esters
  - the individual units are now called nucleotides (have a phosphate)
  - forms a long strand of linked units
  - backbones of sugars and phosphates; one base per sugar
  - nucleic acid encodes info in its structure; info is stored as the sequence of bases
  - nucleic acids can be single-stranded or double-stranded
    - double stranded molecules are more stabilized for long-term info storage
      - thus DNA is ds while RNA is usually ss bc it's a temporary storage medium
    - second strand provides an easy way to replicate or read the info
    - error checking; 2nd strand gives a proofreading system
  - polymerases copy nucleic acids; use one strand as a template to make another strand
  - nucleotides are added one at a time, matching each added base against its compliment on the other strand



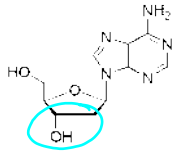
- can use rational drug design; use knowledge of enzyme mechanism and substrate(s) to design a drug
  - the polymerase mechanism uses a nucleophilic displacement reaction
  - we know what the substrates (nucleotides), template strand, and growing strand look like in detail
- selectivity problem with nucleic acids; must block viral enzyme without blocking host enzyme
  - both virally infected cells and normal host cells use the same nucleotide substrate
  - we want to stop the viral nucleic acid production without stopping host nucleic acid production
  - poor selectivity = toxicity; interference w normal cell function (host proteins not being translated)
- non-natural base
  - use a molecule with the same pattern of h-bonding to compliment template strand
  - if the viral polymerase is sloppy (?) the molecule will be integrated into the growing chain and the substrate structure is disrupted; the polymerase cant function
  - chose X so that it is similar enough to get incorporated into the nucleic acid, but different enough that it will make polymerase unable to continue the chain
  - even if the polymerase continues the chain and adds bases after the drug is in the chain, the DNA strand now how ATCG and X in its structure, the structure is unreadable by other enzymes (like ribosome) and thus prevents replication of virus

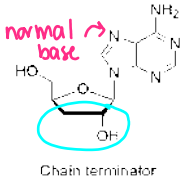


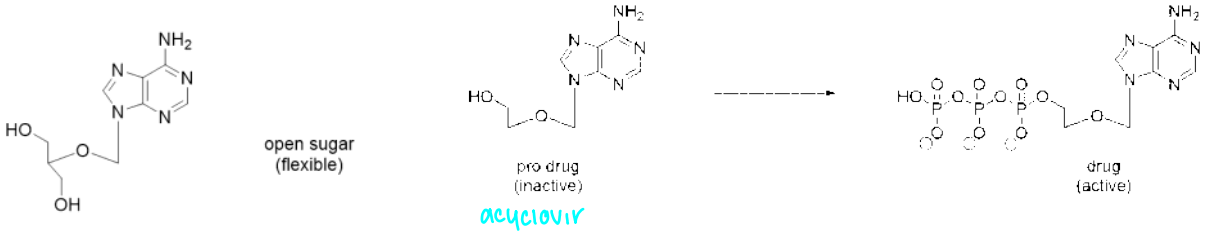
  - requirements for non-natural base:
    - ◇ drug is a substrate for kinases (adds phosphates); bc drug will be phosphorylated in body
      - drug normally doesn't have phosphates on it
      - done by host kinases or virus kinases (rarely)
    - ◇ drug is a substrate for viral polymerase
      - gets incorporated into viral nucleic acid
      - create an "unreadable" strand
        - unreadable by viral pol., host ribosome, etc
    - ◇ drug must NOT be a substrate for host polymerase; bc side effects
      - drug will be incorporated into normal nucleic acids; makes it defective and will lead to cell death
  - examples of non-natural bases:
    - ◇ idoxuridine; introduced in 60s
      - substrate for viral and host polymerase; thus its highly toxic
      - topical use only; this limits exposure to body
        - limited to eye infections of herpes
      - not v successful drug but served as a proof of principle-antiviral drug is possible



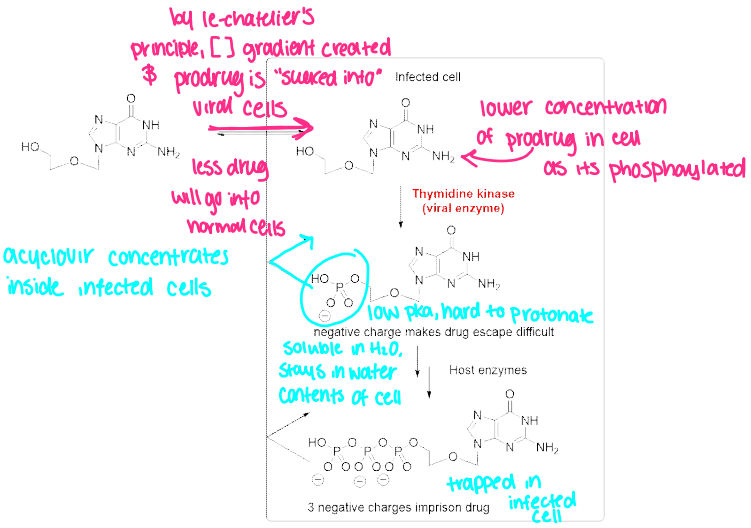
- chain termination
  - use a molecule w a normal base that can be incorporated into chain, but with a non-nucleophilic hydroxy isostere
  - similar enough in structure that polymerase will think there is an OH there and incorporate it into strand
  - stops replication
  - can combine this strategy with non-natural base; achieves selectivity, if you know the X is not recognizable by host pol but is recognizable by viral pol
  - requirements for chain termination:
    - drug is phosphorylated (host or virus)
      - similar enough to a normal nucleoside so its accepted by kinase
    - drug is a substrate for viral polymerase
      - gets incorporated into viral nucleic acid
      - stops replication, get a short piece of nucleic acid
      - viral proteins aren't made
    - drug must NOT be a substrate for host polymerase; side effects
      - this is where the non-natural base comes in; drug should be different enough so host pol doesn't recognize it
  - selectivity is a problem with chain termination
    - drugs have structure that is too similar to normal substrate
      - bc use a normal base
    - occasionally, these drugs get incorporated into host nucleic acid by the host pol
    - many chain terminators are toxic & carcinogenic
  - can improve selectivity by changing the sugar
    - the normal nucleoside is a rigid sugar
    - decided to make a more flexible sugar, lots of bonds that can rotate
    - has the right OH groups in the right distances so hopefully enzyme will incorporate it into chain
    - bc flexible, maybe will be in the wrong orientation to continue synthesis of chain

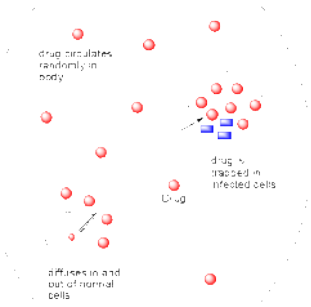






  - acyclovir is a chain terminator drug
    - very high selectivity for virally infected cells
    - low incidence of side effects and low severity of side effects
    - selectivity is primary due to selective bioavailability; its selectively attracted to virally infected cells
    - it's a prodrug; the drug is in its inactive form
    - the drug has to be phosphorylated
    - herpes has a viral kinase: thymidine kinase which can add one phosphate group onto the acyclovir
    - so the drug goes from the prodrug form into the phosphorylated form in the virally infected cell
    - our host enzymes cant phosphorylate the prodrug, so it doesn't become activated in host cells
    - the active form of drug is only in infected cells
    - when the first phosphate is added by thymidine kinase, human kinases add the other two phosphates
    - acyclovir can diffuse in and out of normal cells and stays in inactive form; has no charge so easy diffusion



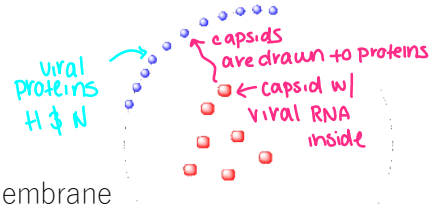


- thus, acyclovir is a v clean drug; accumulates only in virally infected cells
- [drug] in normal cells is too low to cause problems so there are few side effects
- acyclovir was invented by Gertrude Elion aka the drug hunter
  - hard time getting into chem research bc she was a woman, but volunteered to a chem lab and eventually got her masters in 1941
  - war=shortage of chemists so got a job as a food chem analyst
  - in 1944 became a medicinal chemist
  - first woman to head the R&D department of a major pharm company
  - she discovered sooo many drugs (usually v rare to discover even one in ur career)
    - first chemotherapy for leukemia (purenithol)
    - first immunosuppressant
    - drugs for grout, malaria, herpes (Acyclovir)
    - AZT (first AIDS drug)
  - nobel prize in medicine in 1988

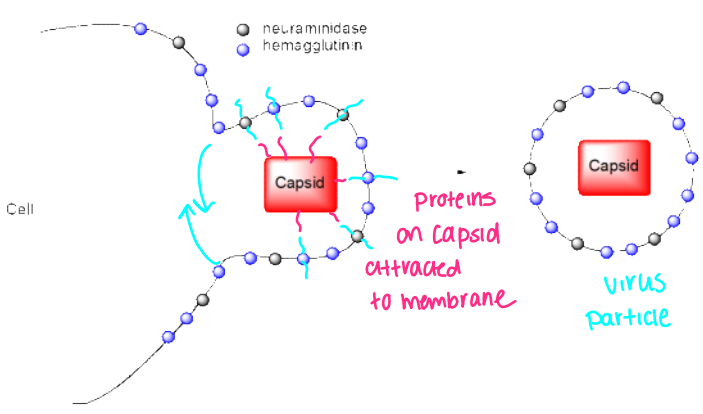
INFLUENZA

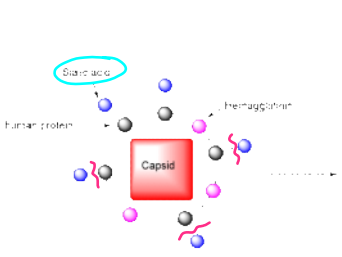
- seasonal respiratory infection
- high mutation rate
  - virus changes yearly
  - virus infects humans and animals; picks up things in one species that can be used in other species; transmission from animals to humans
- most forms not dangerous; elderly (immune system breaking down) and very young children (not fully functional immune system) are exceptions
- occasional pandemics; very virulent strain occasionally arises
  - 1918; 20 million deaths
  - 1957; 1 million deaths
  - 1968 700k deaths
- on average, influenza kills 250k-500k people in the world; 4k-8k in canada per year
- sub types of the flu
  - type A
    - most dangerous
    - pigs, horses, seals, whales, birds, humans
    - lots of genetic variation; hard for us, need new vaccines
  - type B
    - only humans
    - little genetic variation
  - type C; doesn't cause serious diseases

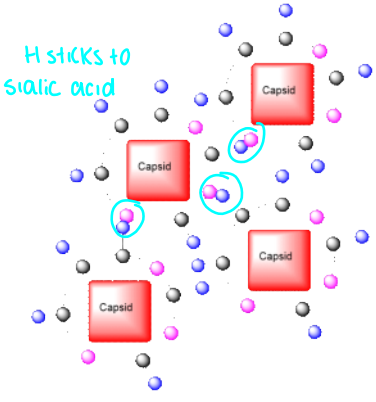
- virus structure
  - enveloped virus; piece of human membrane that it steals
  - envelope is embedded with two viral proteins
    - hemagglutinin (H); impo for viral entry into cell
    - neuraminidase (N); impo for viral maturation (exit from cell)
- bc the two proteins are on the outside of the viral surface, we can detect them in the lab
- flu viruses are classified by serotype (antibody reaction to the envelope proteins)
  - do antibody test to see what kind of antibodies react to the envelope proteins
- there are at least 16 types of H and 9 types of N that we can test for (react to antibodies)
  - hemagglutinins; 16 types named H1 to H16
  - neuraminidases; 9 types named N1 to N9
- most common human flus
  - N1 or N2
  - H1, H2, or H3
  - the other types don't seem to spread well bet humans
- virus life cycle
  - virus hemagglutinin binds to host glycoproteins that contain **sialic acid** (a type of sugar)
  - viral protein interaction w human protein induces formation of an endosome
  - virus envelope and cell membrane fuse; contents spill into cell
  - when in the cell, pH changes on inside of endosome causes it to open and release virus capsid
  - when capsid opens, viral RNA is released and starts to be copied
  - viral RNA is used by ribosome to make viral proteins
  - the replicated viral RNA and viral protein self-assemble
    - capsids form inside cell complete with RNA
    - viral envelope proteins (H&N) accumulate on cell membrane
    - capsids are attracted to surface protein; there is an interaction bet membrane proteins and proteins on capsid that causes the final assembly of virus particle



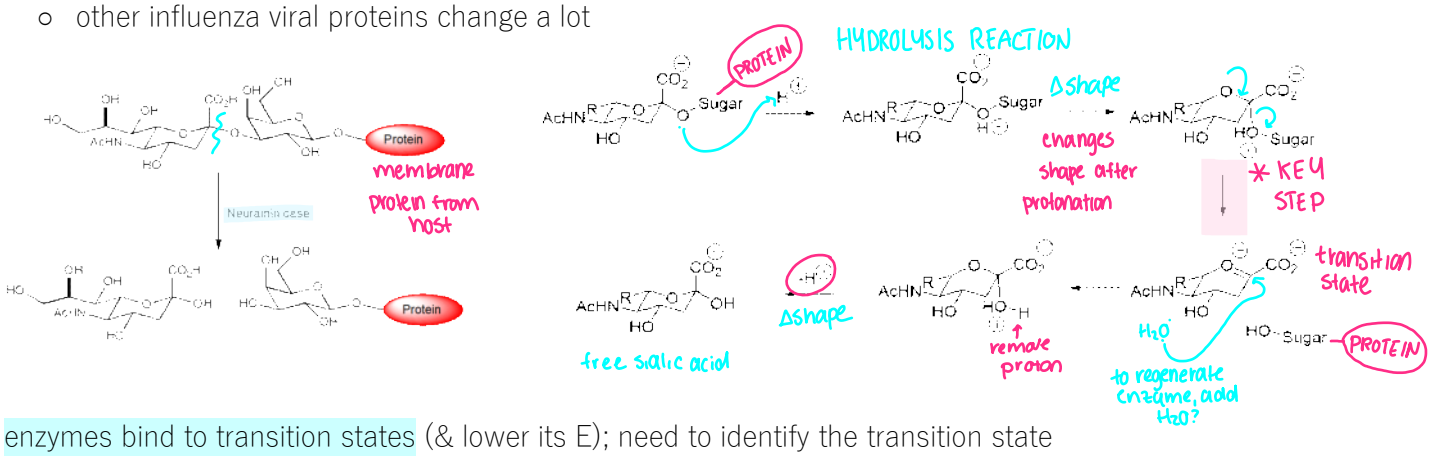
  - virus particles bud from the cell
    - viral proteins span the membrane (inside and outside)
    - inside proteins can recognize viral proteins on the outside of capsid
      - complementary shape & intermolec forces to fit e/o
    - so the membrane wraps itself around capsid and particle is released from cell
- virus particles still contain host proteins (on piece of human membrane it stole)
    - host proteins may contain sialic acid
    - if sialic acid remains on the virus, hemagglutinin from other virus particles will stick to it
    - this will cause the virus particles to clump together and not be able to infect other cells
    - to prevent this, neuraminidase on the membrane removes sialic acid from host proteins on the virus envelope



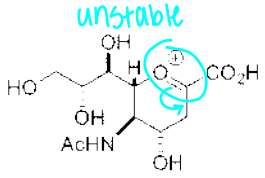




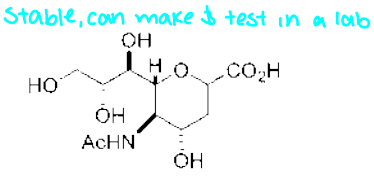
- neuraminidase is an essential enzyme for flu viruses
  - it's a highly conserved protein; meaning it doesn't change its structure much
  - other influenza viral proteins change a lot



- enzymes bind to transition states (& lower its E); need to identify the transition state
- use rational drug design to design a molecule that looks like the transition state of the rxn
  - the drug is called a transition state mimic or transition state inhibitor
  - will fit into active site pocket of enzyme
  - enzyme will think it's a transition state and bind tightly to the drug
  - found that it did inhibit neuraminidase and in cell culture it could inhibit replication of the virus
  - but it wasn't a great drug, needed high concentrations of it, so need something better using this starting point

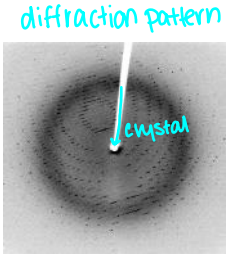


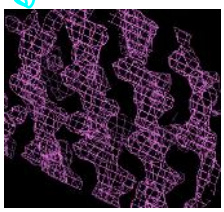
transition state



transition state mimic (inhibitor)

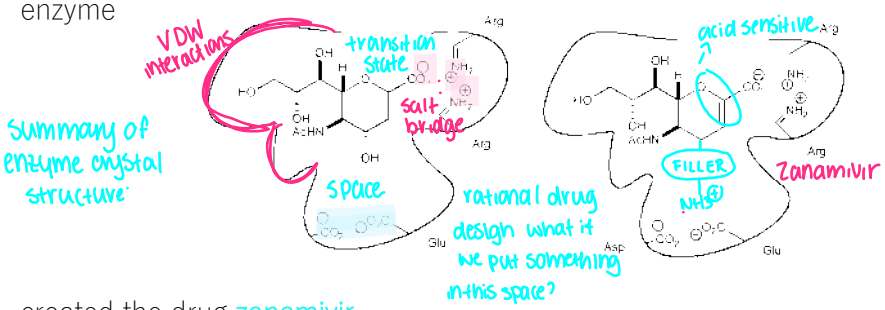
- used an enzyme crystal structure; got it using XRAY crystallography
  - crystallize molecule
  - bombard crystal with coherent (all XRAYs in same direction) XRAY beam
    - wavelength of XRAYs is similar to bond lengths in the crystal (a C-C bond is ~1-3 angstroms=wavelength of XRAY)
    - causes XRAYs to be diffracted
  - XRAYs are scattered by interactions with atoms in the crystal
    - angle det by wavelength and bond distance
    - angle of diffraction can be used to det orientation of atoms in the crystal
    - use the bragg equation (2lamda=2dsintheta)
  - crystals generate large-scale order
    - all molecules are arranged in a fixed lattice
    - this magnifies the effect (all molecules diffract the beam the same way)
    - results in a diffraction pattern
    - can make a map of what the molecule looks like
  - computer can calc the electron density map; a hollow structure showing where atoms are
    - molecular formula usually known or approx known
    - fit the molecular structure into the EDM
  - ribbon diagrams are obtained this way
  - this technique is v impo, used in design of particle accelerators
    - particles accelerate in a circle and give off XRAY when they change direction
    - XRAYs are all coherent, in parallel to e/o
    - prof's old pharm company sponsored particle accelerator manufacture





- bc want to incorporate beam lines which give very HQ XRAY structures

- mixed enzyme w drug inhibitor, allowed inhibitor to stick to enzyme, and crystallized this
- got a picture of the bound structure of the enzyme; can use this to look at all of the close contacts bet drug and enzyme

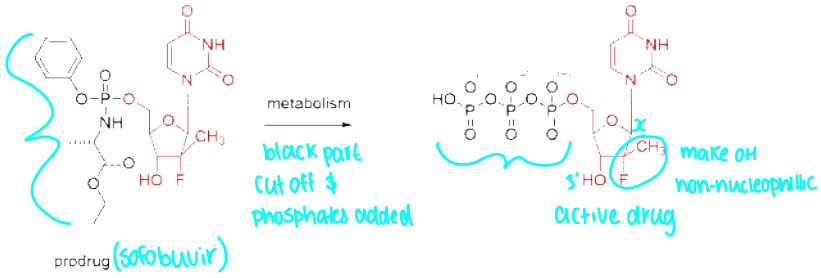


- created the drug zanamivir
  - mild antiviral effect, worked on animals
  - had to be injected, not orally bioavailable bc acid sensitive & violates rule of 5
  - drug was administered as a nasal spray in humans, not that successful
  - was a proof of principle
- came up w a new lead structure from zanamivir (enzyme inhibitor was the first lead)
- after some optimization, came up with the final drug tamiflu
  - 1966-HTS gave no hits
  - 1983-XRAY structure (10yrs to next step bc lots of stereocenters, hard to work w molecule)
  - 1993-zanamivir; proof of principle (in humans)
  - 1997-tamiflu first synthesizes
  - 1999-enerted market
- tamiflu is a very poor drug
  - very expensive bc limited supply (manufacture)
    - ◇ made by semi-synthesis and beginning molecule is hard to get access to
  - must be given in first 24-48 hrs
    - ◇ diagnosis issue; has to be diagnosed v fast
    - ◇ only works for certain subtypes influenza (20% of colds)
  - only reduces course of disease by approx. only 1 day if taken on time
  - FDA benefit/risk analysis indicates clinical use is questionable
  - toxicity issues
  - it's a wide spread drug but only bc of 2009 flu pandemic (ppl are scared)
  - the patent of this drug expired in 2017, so generic versions were made but price didn't lower bc ppl are so scared of flu and will take this drug even though its benefits are so questioned

HEPATITIS C

- causes chronic liver infections, cirrhosis, and can lead to liver cancer
- discovered in 1990
- infects ~2% of US population
- major route of transmission is illegal intravenous drugs (>60%)
  - about 15% of cases are sexually transmitted
  - virus cant survive outside human body
- single stranded short RNA; one protein/peptide is expressed
- the protein is cut into smaller proteins
  - first few are released by host proteases
  - remaining (the majority) are cut out by viral proteases (drug target)
- virus has an RNA dependent RNA polymerase; copies viral RNA (drug target)
  - uses ssRNA to make ssRNA
- virus proteins spontaneously assemble around viral RNA; completes its life cycle
- nobel prize in medicine for 2020 was awarded for hep c discovery

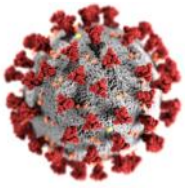
- virus is closely associated with lipids; its viral proteins tend to have lipids form as its outside structure
  - the virus has non-polar nature
  - transported in blood to liver as a complex with its lipoproteins
  - virus is replicated inside membrane of the endoplasmic reticulum (nonpolar enviro)
- viral enzyme drug targets
  - RNA dependent RNA pol (most successful drugs)
  - protease
- the virus has small genome but its complex its proteins do multiple jobs (from video??)
- sofobuvir; a prodrug, the active drug is a chain terminator
  - inhibits viral polymerase and block s replication of viral RNA
  - non nucleophilic isostere of 3'OH by restricting the electron donating ability of the 3'OH group
  - added electron withdrawing fluorine
  - added methyl group that creates sterics, this makes the 3'OH a big nucleophile
    - this reduces electron donating ability (smaller nucleophiles are better)



- ledipasvir; a big molecule
  - hep c virus contains an enzyme which is formed from more than one proteins (it has a quaternary structure)
    - a large protein must bind to a small protein to form a functional viral enzyme
  - the drug sticks to the small protein and prevents protein association and formation of the viral enzyme
  - prevents virus from replication; the enzyme is also involved in replication of viral RNA
- combining sofobuvir and ledipasvir makes the drug harvoni
  - approved in 2014
  - the only small molecule drug capable of curing a viral infection
  - cures HCV infection in 8-24 weeks

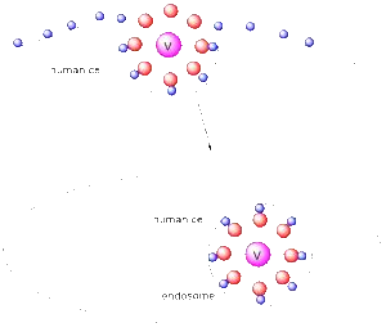
COVID-19

- image formed using cryogenic electron microscopy; use electrons as visualizing tool at very low temps
  - various viral proteins on surface; red protein is the one that helps entry into human cell
- coronavirus family (7 types)
  - most (4) cause colds (mild illness); thus, very little research done before 2002
  - a few cause serious infections
    - severe acute respiratory syndrome SARS (2002)
    - middle east respiratory syndrome MERS (2016)
    - didn't go far bc they involved viruses transmitted animals to humans but didn't transmit well bet humans
  - infect many species; humans, cows, camels, bats, cats, dogs, birds
  - viruses normally don't cross species, very rare
    - cause more severe infection when this happens
    - usually don't see human-to-human when this happens (super rare)



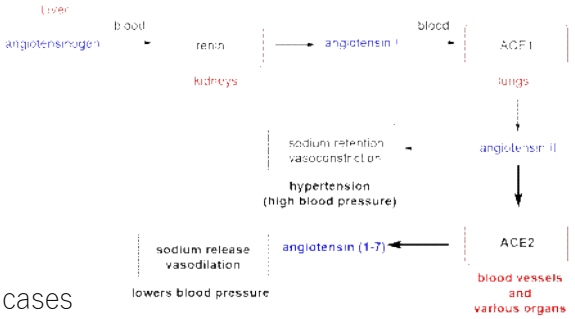
- covid-19 outbreak started in december
- virus is NOT like 4 commonly circulating types (colds); structure is more similar to SARS than colds
- coronavirus (pre 2002)-seasonal colds
  - infected upper airway only (nose and mouth)
  - immune system creates mild response

- SARS, MERS, COVID-serious illness
  - infects lungs and lower airways
  - immune system creates strong response (over-response, reason why ppl are experiencing strong illness)
  - virus damages internal organs (heart, arteries etc)
    - can bind to a protein found in a bunch of organs in the body
- most covid cases are mild; asymptomatic carries
  - no treatment required, self isolate 14 days
- small proportion of population (10%-15%; based on # of positive tests) experience severe illness
  - most older than 60 and have pre-existing medical condition (any age):
    - diabetes type II
    - heart disease (heart attack, CAD, cardiomyopathies (inflammation of heart))
    - hypertension? ambiguous, on some lists and not on others
    - COPD; chronic obstructive pulmonary disease
    - chronic kidney disease
    - cancer
    - immunocompromised state
    - sickle cell disease
    - obesity
  - other conditions might increase risk
- older ppl experiencing severer disease (more older ppl being hospitalized)
- has not been time to develop drugs designed specifically for covid; recall antiviral drugs are v specific to each virus
- existing antivirals are tested for activity against covid; commercially available drugs developed for other viruses
  - tested in lab (invitro) and quick animal test, then test in humans
- failed antiviral drugs tested for activity against covid
  - development re-started; may or may not be commercially available
- virus life cycle
  - entry into cell via endosome
  - translation of viral RNA (ssRNA) and immediately expressed using host ribosome
  - processing of viral proteins (2 viral enzymes; drug targets)
  - viral RNA is copied (2 viral enzymes; drug targets)
  - virus capsid assembly (including RNA)
  - virus envelopment w host membrane and virus release by exocytosis
- viral entry into cell is mediated by binding to ACE2 receptor
  - actually an enzyme that serves as the receptor for covid
  - ACE2 is impo in regulation of blood pressure etc, normally acts as an enzyme on certain cells of body
  - when virus sticks to ACE, it's the protein that is used to get into cell
  - bc there is binding bet proteins, its designated (renamed) as a receptor
  - the red spike protein outside covid molecule is what binds to ACE2
  - virus wraps some membrane around itself and proteins stick to e/o and eventually form endosome



- from list of pre-existing medical conditions above ^^, all except immunocompromised state are linked to ACE2, so binding of virus to ACE2 could explain the effects of covid on ppl w these diseases
- ACE2 is part of the renin-angiotensin system in the body
  - liver produces angiotensinogen (hormone)
  - renin (an enzyme) in the kidneys processes angiotensinogen and cuts it into pieces
    - one of these pieces is angiotensin I (hormone)
  - angiotensin I circulates in blood until it gets to lungs where there is the enzyme ACE1
  - ACE1 processes angiotensin I into angiotensin II

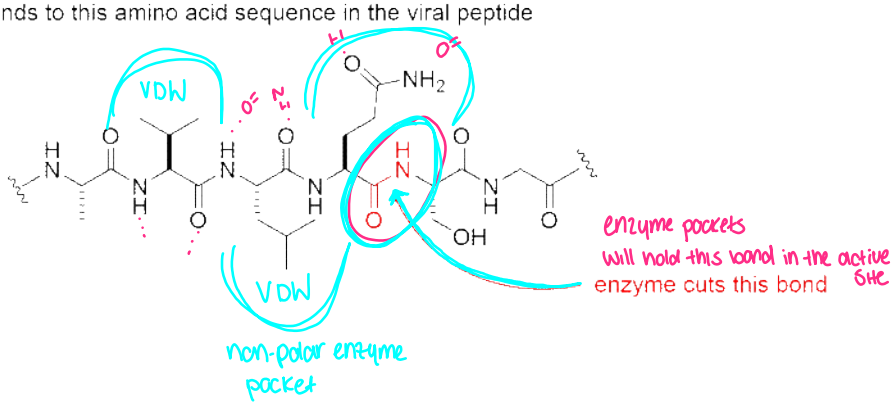
    - angiotensin II raises blood pressure aka hypertension (via sodium retention & vasoconstriction)
        - angiotensin II is also associated with hypertrophy, fibrosis, oxidative stress
    - ACE2 down regulates the system
        - it converts angiotensin II to angiotensin (1-7) which acts to lower blood pressure (via sodium release and vasodilation)



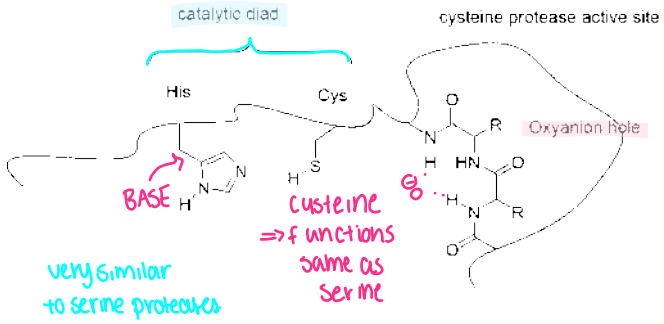
- ACE2 found in arteries, heart, lungs, kidney, intestines, brain
    - lowers BP, and is anti-inflammatory, anti-fibrosis
- binding to covid spike protein deactivates the ACE2 enzyme
    - no catalytic function when bound to spike (can raise BP)
- endosome formation removes some ACE2 from cell surfaces
- increased amounts of ang II, decreased amounts of ang (1-7)
- thus, virus infects entire body, not just lungs
    - so we see complex inflammation and organ damage in severe cases
        - acute lung injury, blood clots in arteries, edema in heart in brain, pancreas damage, & liver damage
- two drugs hydroxychloroquine and chloroquine; usually used to treat malaria
    - bases (raise pH); two nitrogens on the molecule are bases
    - when covid is in the cell, the pH lowers in the endosome which causes it to uncoat and open, releasing the viral contents
        - these drugs can maybe prevent endosome from opening
    - the drugs were used a lab reagents in 90s to study endosome formation of viruses
        - saw anti-viral activity by acting as a base (using large concentrations, 60-300uM)
    - works at high doses in cells (30uM) and mice
    - used in humans for other viruses; generally need high doses
    - use on covid:
        - first report of clinical benefit (humans) in february
            - <100 patients; no data included to judge whether it was good or not
        - first clinical trial in february
            - 20 patients, non-randomized, external controls
            - saw reduced viral loads w 600mg/day doses
        - larger clinical trials since then have seen little to no benefit
        - bc of this, FDA withdrew their approval of these drugs for use in covid
            - bc benefits don't outweigh risks (there are side-effects with the drugs)
- translation of viral RNA forms a long polypeptide containing all viral proteins (polyprotein=viral peptide)
- two viral proteases that cut out viral proteins which then fold into their functional shape
    - 3CL protease
    - papain-like protease
- how does the enzyme know which bonds to cut? recognizes a particular amino acid sequence in the viral peptide
    - enzyme shape and non-covalent interactions matches the substrate shape
    - the pockets will line up and hold the amide bond that gets cut in the right place (in the active site)

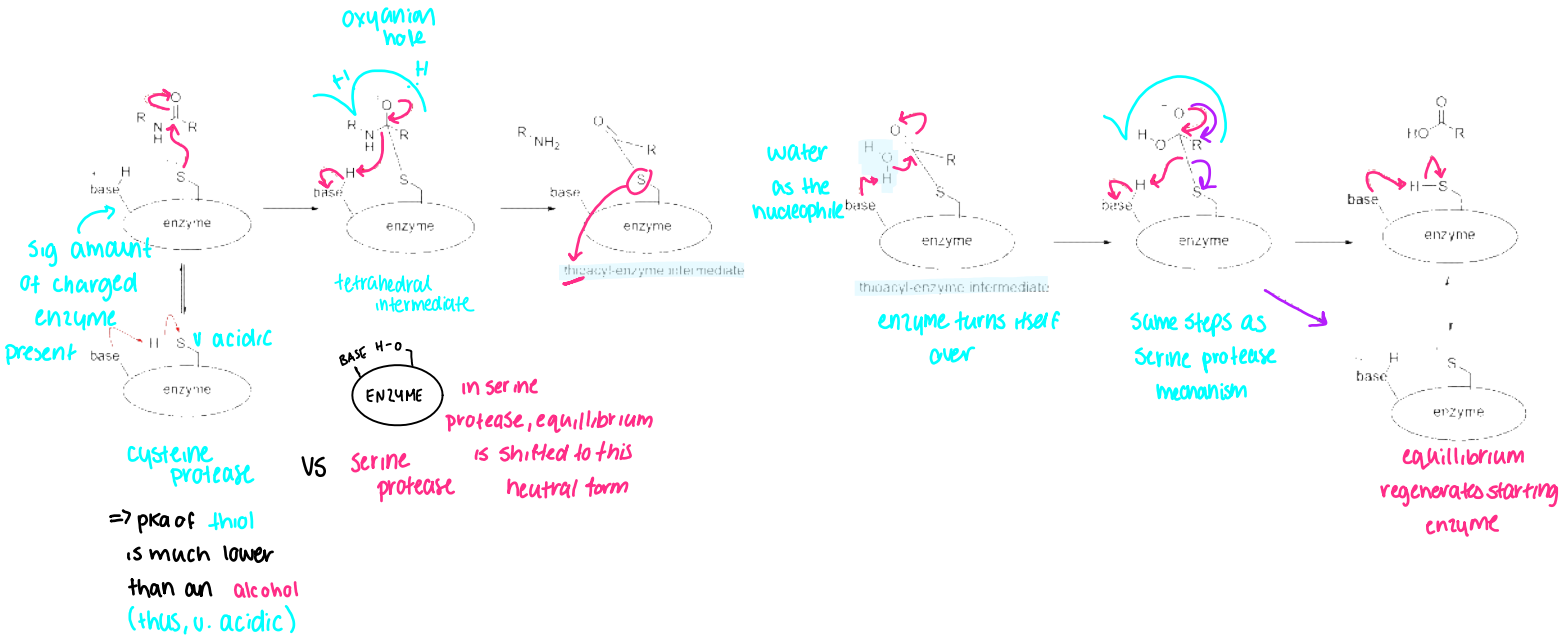
each protein has the sequence @ its c-terminus

3CL protease binds to this amino acid sequence in the viral peptide

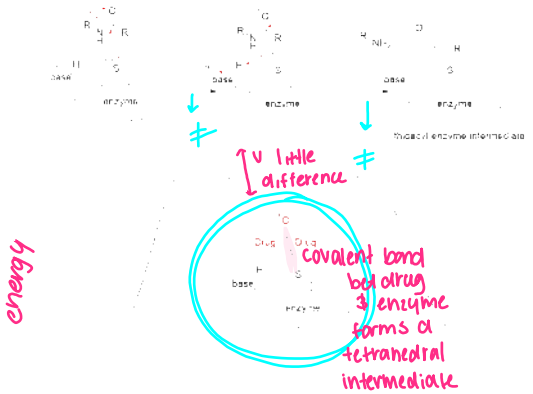


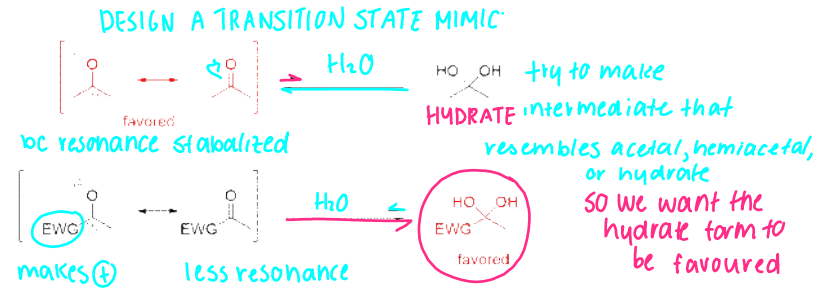
3CL protease is a cysteine protease:

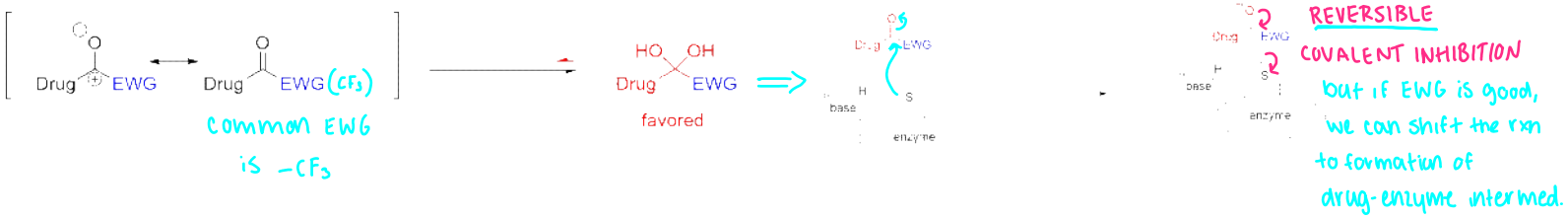


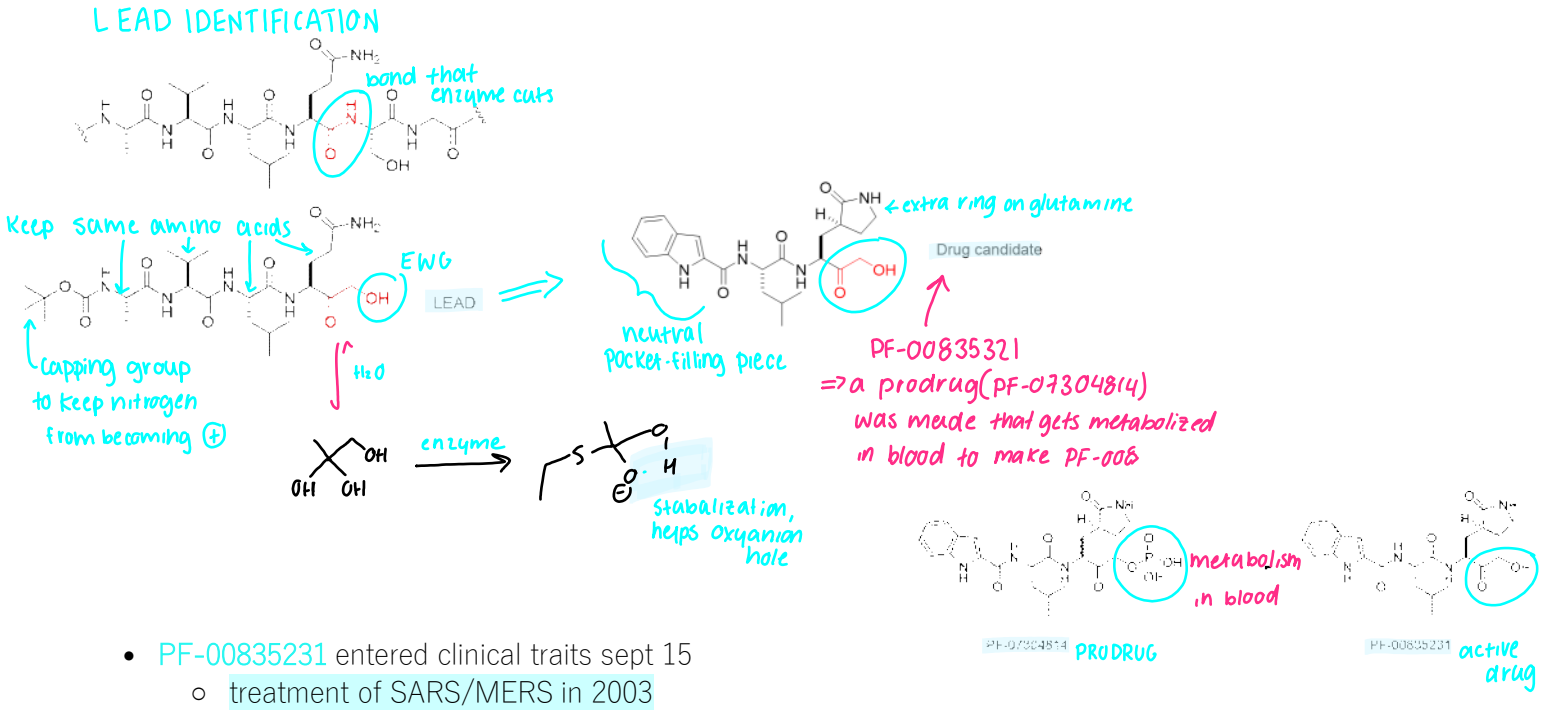


rational drug design: which molecule is this mechanism does the enzyme bind the tightest to (transition state)?

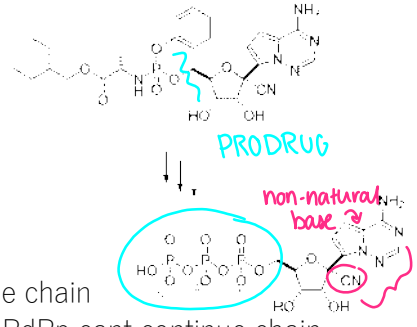








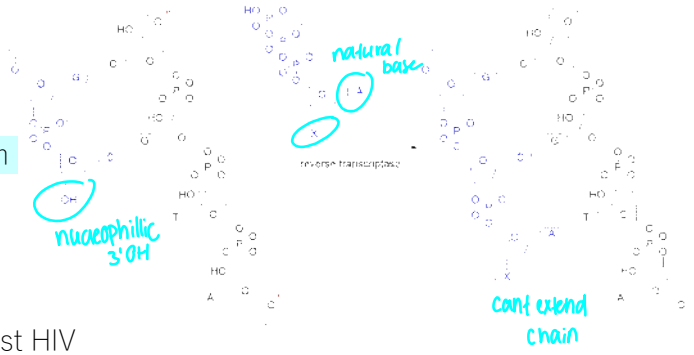
- PF-00835231 entered clinical traits sept 15
  - treatment of SARS/MERS in 2003
    - development halted bc outbreaks contained (no market)
    - development re-started for covid outbreak
  - prodrug; metabolized to PF-00835231 in blood
- after viral RNA gets copied and makes viral proteins, the viral proteins causes viral enzymes to copy the viral RNA that goes into capsids
  - 2 enzymes used to copy viral RNA
    - RNA dep RNA pol (RdRp); copies ssRNA (nucleic acids)
    - helicase; unwinds nucleic acid
- can target viral polymerase (RdRp) by using a drug that looks like a nucleoside/nucleotide
- remdesivir (a re-purposed ebola drug); a nucleotide drug approved for use on covid
  - a prodrug that gets metabolized; left part cut off and addition of 3 phosphate groups; makes an artificial nucleotide that acts against polymerase
  - has a non-natural base and cyano group (CN) that work together to give the mechanism of operation
  - developed in 2009 for hep c by same company that developed harvoni
    - failed bc not efficient enough
  - tested against ebola when there was an outbreak
    - saw in-vitro activity
    - clinical trials during 2015 outbreak
    - early success, but ultimately failed in long term clinical trials (2018)
  - chain terminator of covid
    - gets put into replicating strand by RdRp
    - but bc it has a non-natural base, an abnormal shape is created and the chain continues for a bit (1 or 2 more bases) before the shape is wrong and RdRp cant continue chain
      - a delayed-action chain terminator
  - reduced recovery times from covid in some studies
  - must be used early in the infection
    - animal studies only show improvement if administered within 24hrs of infection
  - FDA authorized emergency use in spring
- hasn't been enough time to come up w covid specific drugs yet, scan existing drugs
  - protease inhibitor is pretty close target of covid though



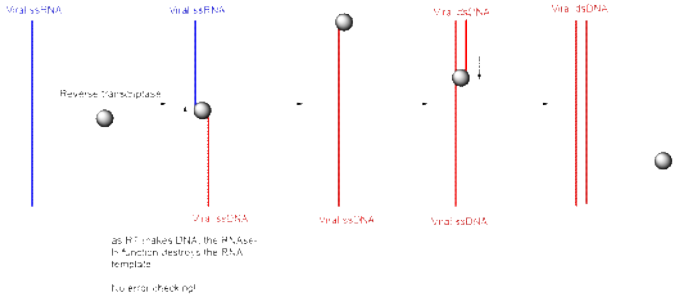
HIV

- first cases diagnosed in 1981
  - immune collapse
  - kaposi's sarcoma (rare disease, skin cancer)
- patients suffered from T-cell depletion (key component of the immune system)
- victims die from opportunistic infections that are normally controlled by a healthy immune system
- HIV (human immunodeficiency virus)=agent that causes AIDS
  - AIDS is the later stage of HIV infection when person suffers from opportunistic infections
  - discovered in 1984
- spread through exchange of bodily fluids
  - through unsafe sexual practice
  - intravenous drug use
  - congenital infection of newborns during delivery (less common)
- targets CD4 cells (a type of T-cell) of immune system
  - CD4 is a protein on the surface of the cell; important component of the immune system
  - virus has a protein that sticks to CD4
- origin of AIDS
  - oldest known sample from 1959; blood sample from the belgian congo
  - estimated that virus originated 1920s
  - virus may have spread to humans from chimpanzees b4 1920s
- huge outbreak in mid 1990s; became the world's most deadly infectious disease
  - 0.8% of all adults HIV positive
  - leading cause of death in africa; fourth leading cause of death in the world (over 1million deaths/yr)
- HIV today is more manageable bc developed world
  - still the 4th leading cause of death in low-income countries
- untreated HIV has 3 phases:
  - acute phase; lasts 6-10 weeks
  - chronic phase; lasts 8-12 years if untreated (shorter if person's health is bad)
  - AIDS phase; 2-4 years if untreated (show symptoms, ends with death)
- to treat, can keep person in a long-term chronic phase (cant get rid of virus)
- the acute phase consists of flu-like symptoms
  - there are high levels of virus
  - phase ends when immune system reduces viral loads (almost gets it under control)
- during the chronic phase person is asymptomatic but fully infectious (thus this phase is a big reason why AIDS is so widespread)
  - viral replication is v high
  - immune system destroys most viral particles made (>99.999%)
  - a few virus particles escape to continue infection
  - CD4 cells in immune system are slowly depleted
  - phase ends when immune system runs out of CD4 cells
- the AIDS phase starts when immune system is destroyed
  - lots of secondary infections; these cannot be cleared by the immune system
  - person is fully infectious
  - viral replication is v high
  - person dies of secondary infection
- HIV is a retrovirus; transmit genetic info in a reverse direction
  - normally: DNA to RNA to protein
  - retrovirus: ssRNA to dsDNA
  - viral genome is carried in a ssRNA

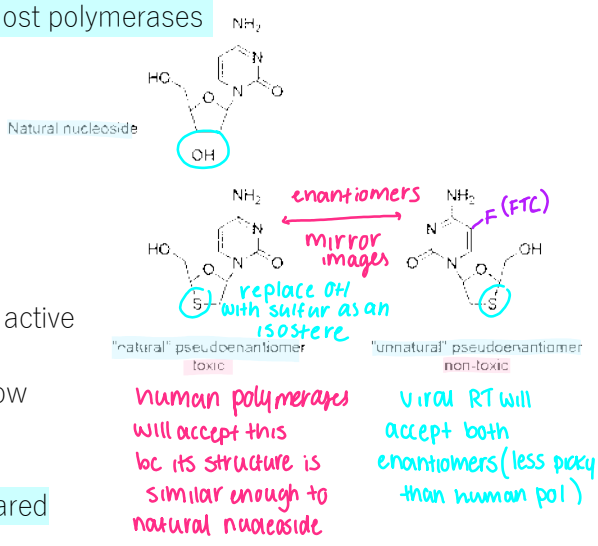
    - viral enzyme called reverse transcriptase copies the ssRNA into dsDNA
        - reverse transcriptase has two catalytic domains:
            - one functions as a polymerase
            - the H domain destroys RNA template as DNA gets made
                - impo for survival of retrovirus and the key reason why immune system cant clear the virus
    - reverse transcriptase destroys the RNA template (RNAse H domain) as it is copied
    - reverse transcriptase then uses the ssDNA to prepare ds viral DNA
- HIV replication cycle
    - virus binds to CD4 protein on outside of T-cells
    - virus envelope fuses with host membrane and injects capsid inside
    - capsid opens, RT & viral RNA is released; RT then copies viral RNA into dsDNA
    - viral integrase enzyme inserts viral DNA into host chromosome; infection is now permanent, cell cant get rid of the virus
        - traces of retroviral infections in ancestors are found in our DNA today
    - viral genome "activates", viral RNA is expressed, single viral polyprotein is made
    - viral protease cuts itself out of the viral protein, and cuts the rest of the viral protein into smaller viral proteins
    - virus particles self-assemble around viral RNA
    - virus buds from the cell taking membrane piece with it (envelope)
- HIV viral drug targets
    - 4 enzymes
    - several good targets
        - reverse transcriptase
        - protease
        - GP41; part of the envelope protein of virus (acts like a drill)
    - other targets
        - integrase
        - RNAse H (part of RT)
        - GP120
- can use a chain terminator to stop DNA strand completion
    - use non-nucleophilic hydroxy isostere
    - results in DNA chain not being extended



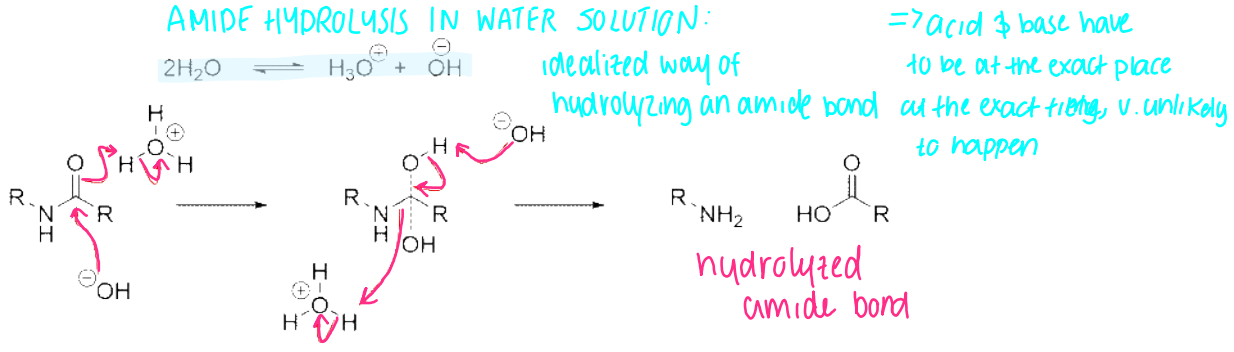
- AZT is a drug discovered in the 60s
    - failed as a cancer drug bc highly toxic
    - based on its mechanism of action, was tested against HIV
    - gained FDA approval after truncated clinical trial
        - trial was supposed to be 10 months but patients improved after 2-3 months
    - later, someone repeated the clinical trial and shows the drug was ineffective
        - short period of antiviral activity
        - then resistant virus appears in a few weeks and drug no longer works
- drug failure due to resistance; virus has high mutation rate
    - RT is error prone bc enzyme destroys the template RNA as its copied
    - no error checking, sloppy replication
    - so every virus particle is different by at least one base pair, v high mutation rate



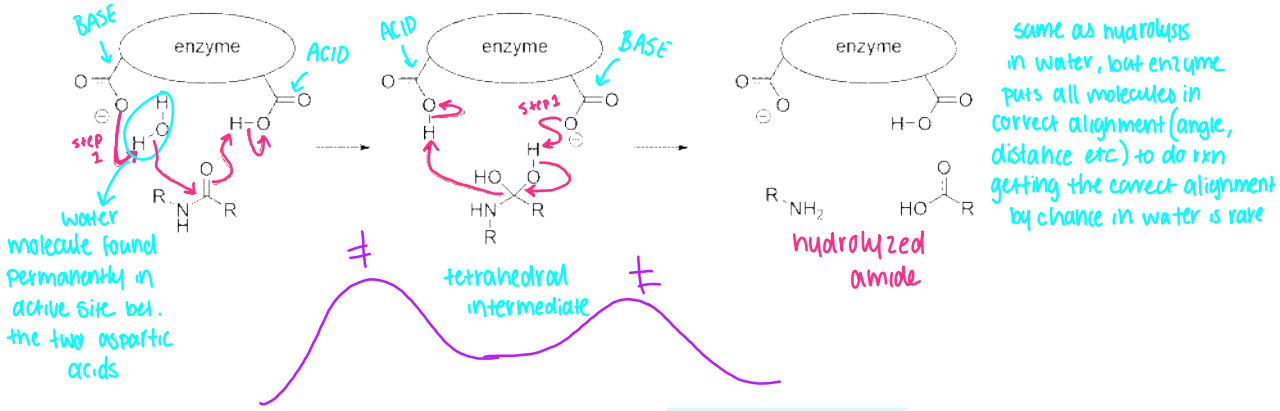
- 3TC racimate
  - drug that was originally made as a racemic mixture; company found that both enantiomers inhibit RT
    - RT was not picky
  - but the natural enantiomer is highly toxic bc it's also substrate for host polymerase
    - the "un-natural" enantiomer hos low toxicity; its not recognized by host enzymes
  - so they got two drugs: pseudoenantiomers of the natural nucleoside with different toxicities
    - "natural" pseudoenantiomer is highly toxic
      - substrate for host polymerase & kinases
    - "un-natural" pseudoenantiomer has low toxicity
      - substrate for host kinases but not recognized by host polymerases
  - drug was sold is the "un-natural" enantiomer; low toxicity
  - 3TC was discovered by bernard belleau in 1989
    - canadian uni prof
  - the 3TC patent was poorly written; let others steal the drug
    - described as a racimate
    - patent described that the isomers could be separated but didn't describe how, patent filed 1989
    - the company later discover that both enantiomers were active
    - filed another patent 1991
    - patented the distinct enantiomers but didn't describe how to make them
  - the company proceeded with clinical trials
  - low toxicity, but effectiveness was limited bc resistance appeared
  - by 2000, drug was selling $1B per year
  - dennis liotta noticed the patent didn't address the non-obvious aspect of how they made the enantiomers
    - he found a way to make the pure enantiomers and patented both the syntheses
    - this gave him ownership of the enantiomers and held rights to their drug and got money from them
  - liotta also noticed that the patent missed a key compound
    - when you patent drugs, you also have to patent many molecules that could be made that fall in the same family of your drug
    - the fluorinated analogue of 3TC (called FTC) was a common change when non-natural base anticancer drugs were developed
    - so he stole their whole patent and created a patent for FTC



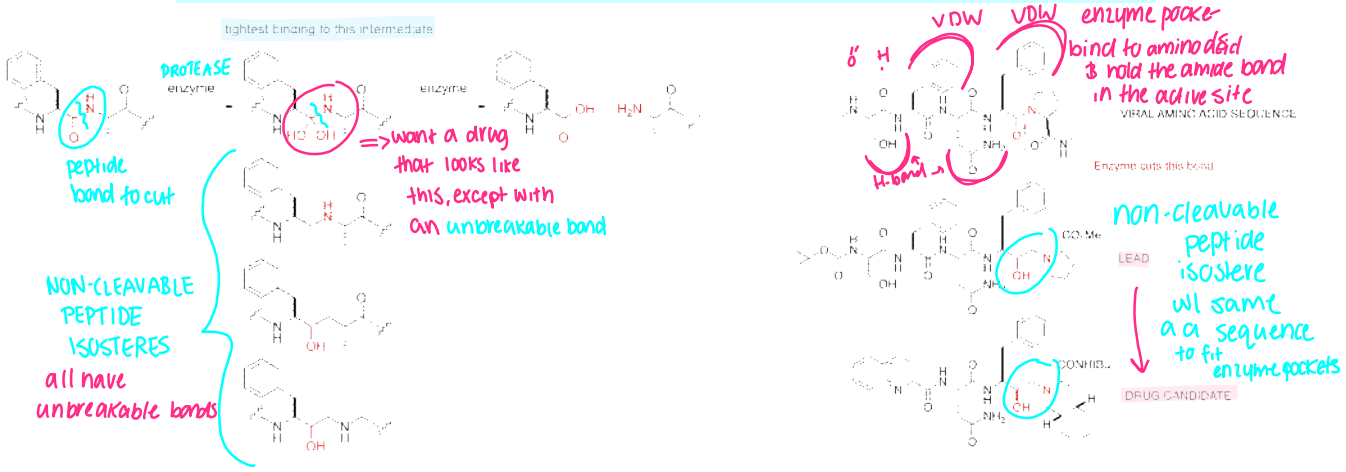
- HIV protease is an essential enzyme; it cleaves the long polypeptide into smaller active viral proteins/peptides
  - member of a family of proteases called aspartyl proteases
    - active site has 2 conserved aspartic acids
    - protease hydrolyzes an amide bond (similar to how it would be hydrolyzed in water solution)
  - aspartyl protease sets up acid and base in the exact location to carry out rxn



AMIDE HYDROLYSIS BY ASPARTYL PROTEASE



- now that we know the mechanism of HIV protease, can do rational drug design
  - the tetrahedral intermediate is v close to transition state; binds tightest to the enzyme



- most of the cost of drugs is not the active ingredient (manufacturing cost is v small)
  - its marketing, research etc
- until HIV/AIDS, no drug program has ever been cancelled bc the drug itself was too hard to manufacture
- protease drugs were v complex and v hard to manufacture
  - large molecules w lots of stereogenic centres
  - acyclic molecules (open chain)
- also had v poor bioavailability (rule of 5) so needed a big dose
- might be impossible to manufacture economically
- so some companies stopped before development; afraid to be forced to continue if the drug was good
- merck proved it could be done: crixivan (indinavir)
  - used brand new catalytic methods to construct stereocenters (got lucky)
- protease drugs also didn't work well bc of resistant forms of HIV
  - HIV replicates rapidly, replication is sloppy
    - genome of each virus is diff, resistant mutants generated each day
  - anti-HIV agents select resistant strains unless all viral proliferation is blocked
    - is ur drug doesn't kill everything, you're basically selecting for the resistant strains bc they are the only ones that survive the drug and replicate
  - no single anti-HIV again is powerful enough to do this
- combination therapy; combine three diff HIV drugs together
  - the chance that a virus particle is resistant to three drugs at the same time is really small
  - if one drug fails, second one may still work, third drug provides insurance
  - result: 50% of patients achieved low viral loads within 6-52 weeks
  - aka highly active antiretroviral therapy (HAART) aka ART
  - combination of 2 nucleoside analogues and one other inhibitor (protease, or integrase, or NNRTI?)
  - resulted in a huge death rate decrease

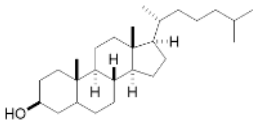
- drawbacks to HAART (1st generation)
    - not v user friendly, compliance issue
        - not v bioavailable so have to take 16+ pills a day
        - complex regimen (with/without food, empty stomach, w/o other drugs etc)
        - side effects: rash, nausea, nightmares, headache, anemia, hepatitis, lipodystrophy
    - compliance problems produce resistance (ppl weren't consistently taking pills, or just stopped)
    - cost ~$10k/year
- so HAART was initially successful bc improved quality of life (no secondary infection)
- but severe side effects and inconvenient dosing; your life was built around your meds
- appearance of resistance and increased speed bc patients skipped meds
- 2nd generation drugs were designed to be more "drug-like"
    - reduced side effects & convenient dosing
- current therapies are based on clean drugs
    - the pseudoenantiomer nucleosides (3TC and FTC)
    - few side effects, not substrates for human enzyme
    - good patient compliance
    - combined with integrase inhibitors or 2nd gen protease inhibitors (improved properties)
- today, HIV is manageable infection in developed world
- disease can be cured, requires continual use of drugs
    - small # of ppl have been cured, req extreme treatment
- better access to ACT now, decline in deaths
    - still an issue in less developed places

TOPIC 6: LIPITOR

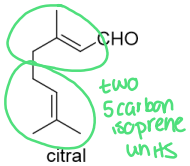
October 27, 2020 12:04 AM

- lipitor=all time best selling drug
  - used to control production of cholesterol in the body
- cholesterol is an essential substance, found in all animals
- important constituent of membranes
  - provides strength & helps create proper liquidity
- continuously made & recycled in the body to maintain appropriate concentration
- "goldilocks" chemical
  - minimum amounts are req for good health but too much causes problems
- high blood cholesterol is associated with heart attacks
  - blocks the arteries (specifically coronary arteries) and stops blood from getting to the heart
- cholesterol is a triterpene (a class of molecules)
  - lipophilic molecule
  - made of 6 isoprene units; fundamental building blocks of terpenes

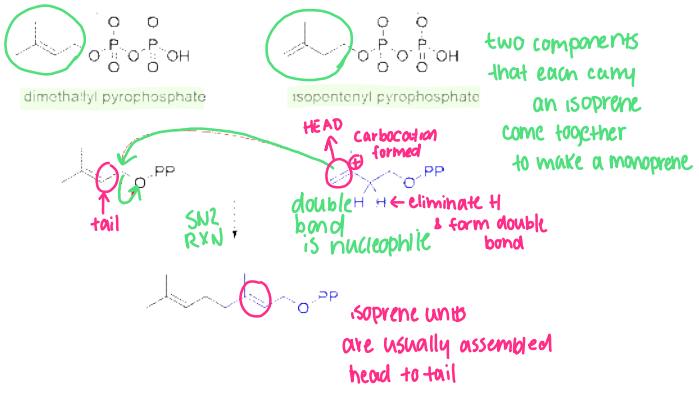
cholesterol



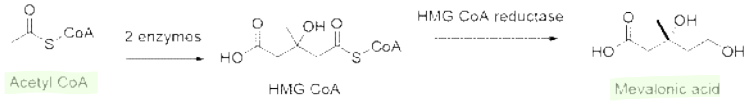
- terpenes; secondary metabolites formed from isoprene units
  - monoterpenes; made from 2 isoprenes (10 carbons)
    - ex: citral
  - diterpenes; made from 4 isoprenes (20 carbons)
  - triterpenes; made from 6 isoprenes (30 carbons)

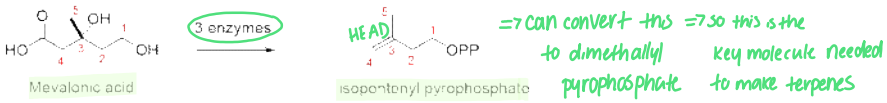


- assembly of isoprene units:

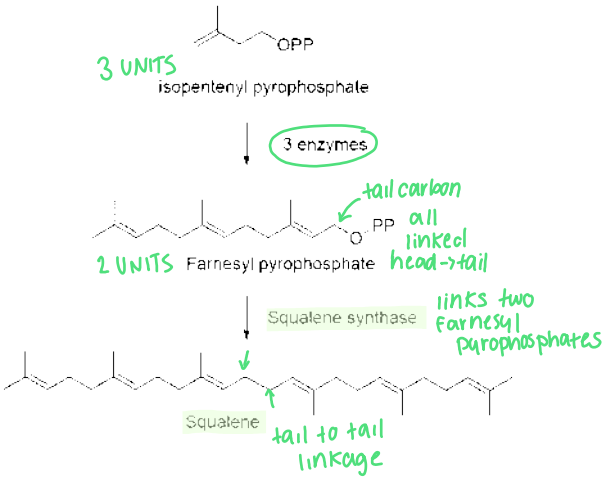


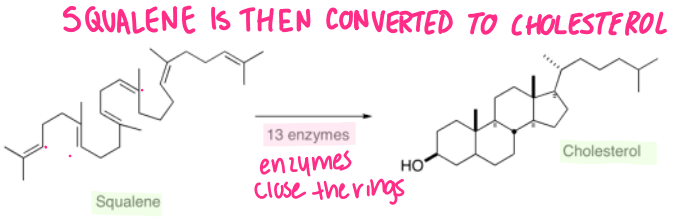
- cholesterol is formed of 6 isoprene units
  - 2 are linked head to tail
  - those 2 units ^^ are linked to e/o tail to tail
  - after this, there is some rearrangement of atoms to make final structure
- where does our cholesterol come from?
  - diet makes a v small contribution
    - we have cholesterol in our diet, but our digestive system is bad at absorbing it
    - food from animal sources
  - majority of our cholesterol is made in liver
    - saturated fats and trans fat in our diet is associated with this
- cholesterol is made in the liver from saturated fat
  - staring from acetyl CoA, 23 chem rxns with 23 diff enzymes used to make cholesterol
- synthesis of mevalonic acid:





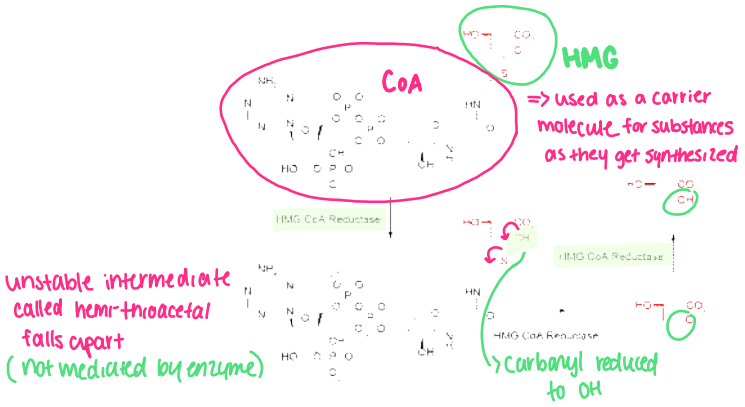
- six units of isopentenyl pyrophosphate form squalene, which is converted to cholesterol:

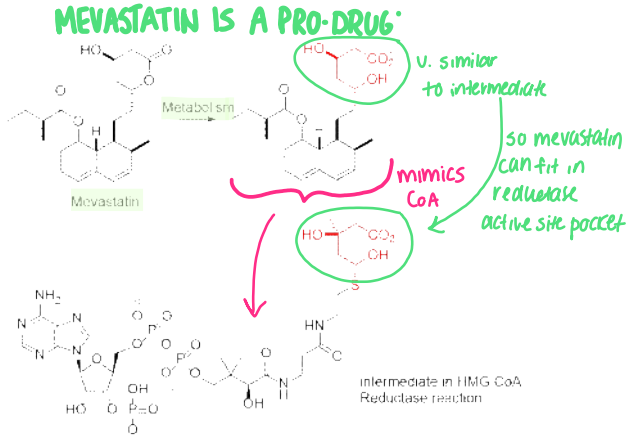




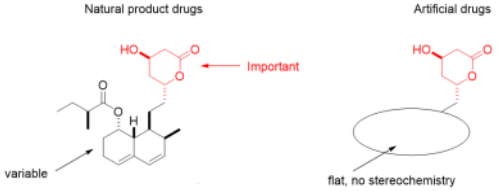
- so cholesterol is biosynthesized in the liver, then distributed to all cells in the body via blood
- bc body makes a big investment into making cholesterol (23rxns & enzymes), it doesn't want to throw it away
  - so excess cholesterol is stored in arterial walls (prof thinks body is stupid for this)
- cholesterol thus can build up inside the arterial lining-plaque
  - diameter of artery (opening) doesn't change as this happens
  - the artery bulges out instead
- cholesterol in the plaque becomes oxidized into cholestenone
  - the body doesn't like this and tries to get rid of cholestenone
  - oxidized cholesterol attracts macrophages
- macrophages consume the cholestenone and become foam cells bc have a lot of lipophilic material inside (look like soap cells)
- the macrophages secrete substances that damage the muscle cells that form the lining of arteries
  - this damage can result in a rupture of artery (a cut in the artery)
  - this is when blood clots form in the artery
- so its not the cholesterol (plaque) that blocks the artery, its blood clots
- to prevent storage of cholesterol in arterial lining, need to prevent production of cholesterol
  - can prevent heart attack due to this by changing diet since cholesterol is produced from saturated and trans fat in diet
    - need to reduce total fat in diet
    - avoid processed foods; lots of saturated fat & trans fat
    - reduced meat intake, eat fruit and veggies
    - reduce total food intake
  - average north american eats 2X required food
    - ppl reluctant to change eating habits, we like food
    - so use drugs to block cholesterol biosynthesis, more user friendly than changing your entire diet
- triparanol (1959); drug effective in lowering blood cholesterol
  - but caused cataracts, removed from market in 1962
  - used as an example in medical textbooks, so physicians were afraid of cholesterol lowering drugs
  - inhibits the last enzyme in cholesterol synthesis, so the other 22 enzymes till make intermediates
  - results in rxn stopping at desmosterol (insoluble) and causing it to build up
  - thus, was a toxic drug
- so want to inhibit something early in the sequence for safety
  - build up of synthetic intermediates can cause toxicity
  - less build up of intermediates if block an early enzyme

- mevastatin discovered by prof in japan in the 70s
  - competitive inhibitor of enzyme called HMG CoA reductase
  - one of the early enzymes in sequence
  - the enzyme works as a reducing agent (like $NaBH_4$ or $LiAlH_4$)

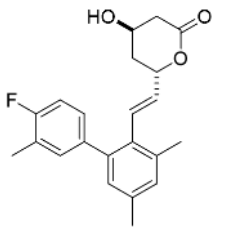




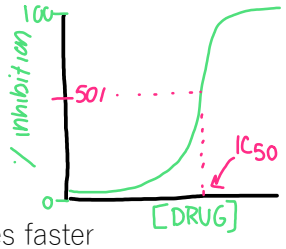
- HMG CoA reductase is 3rd in the sequence
  - If its blocked, not a lot of buildup
  - build up a little material, but they are simple substances the body can use for other things
  - less side effects this way
- merck discovered lovastatin in a species of mushroom; no toxicity issues
- other companies identified naturally occurring substances with similar structure that could inhibit HMG CoA reductase; so there are a few statin drugs that were discovered
- but there was a marketing issue; triparanol made doctors scared to prescribe statins, even though there was evidence they were safe and effective
- so drug companies had to re-educate doctors (bc they don't know the chemistry behind it)
  - triparanol example in textbooks suggested inhibited cholesterol prod=dangerous
    - ◇ no mention that it was bc enzyme was late in sequence and caused build up
  - new drugs inhibit diff enzyme later in sequence avoiding buildup problem
- pharm companies carried out a clinical trial as advertisement to doctors that its safe
  - four-S study of simvastatin; scandinavian simvastatin survival study
    - ◇ 4444 patients
    - ◇ 35% reduction in cholesterol & 42% less likely to die of heart attack
    - ◇ non heart related deaths at normal rates; no bad side effects
  - sales of all statin drugs rose after this study
- drug companies in this business were stuck with natural materials; statins have complex structures
  - its hard to manufacture, need to use natural source to make the drug for you
  - issues: need to make sure you have enough of the source and its hard to make any changes to the molecules (hard to fine tune its structure)
- so companies were looking for synthetic versions with simpler structures so you can tune its properties
  - used rational drug design
  - the HMG part is important, cant change this bc enzyme is picky
  - the CoA mimic is variable; the system of six-membered rings is flat and lipophilic
    - ◇ but has stereochemistry, want to get rid of this bc its so hard to work with
  - so need to make the CoA mimic flat and lipophillic, but with little to no stereochemistry



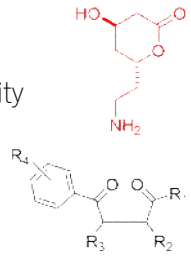
- merck attempted of make an artificial statin but didn't get anything good
  - made a simple ring system; flat, lipophilic, and no stereochemistry
  - they terminated the project and published their results
  - other companies used it as a template to design their own artificial statin

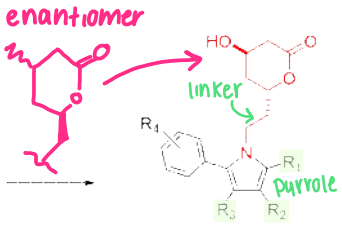


Merck HMG-CoA inhibitor

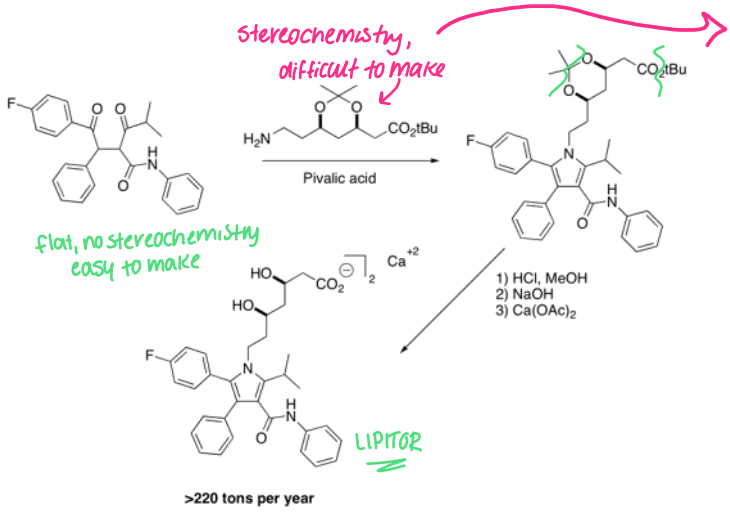


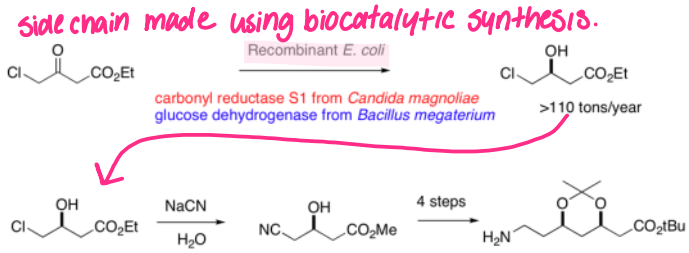
- pfizer designed a pyrrole-based ring system based on merck publication; lipitor
  - merck's has a lot of C-C bonds (aromatic rings/double bond)
    - these are hard to make using synthetic rxns and slows down the work
  - using a pyrrole makes it faster to make a lot of diff molecules and optimize structures faster
  - SAR: linker optimization; tested diff linkers
    - IC_{50}=drug concentration for 50% inhibition of enzyme
      - at 0% inhibition, enzyme is working at full capacity
      - at 100% inhibition, enzyme is working at 0% capacity
    - want the smallest IC_{50} value possible
    - found that a two carbon linker is optimal (IC_{50}=0.5)



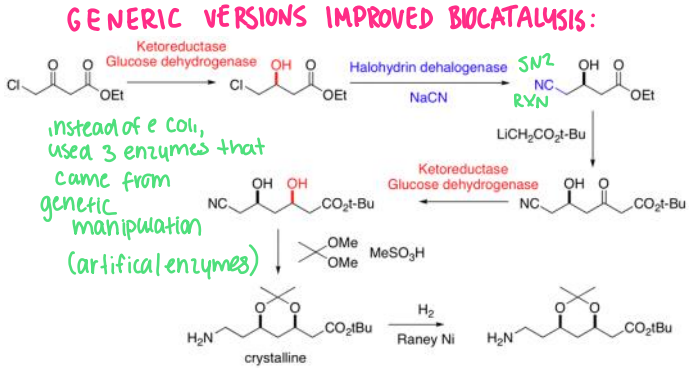


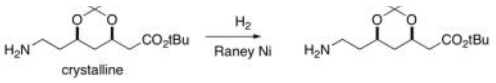
  - SAR: pyrrole 5-substituent at R_1
    - tested 30 compounds' IC_{50}
    - found that an isopropyl is optimal (IC_{50}=0.4)
  - overlay of the molecule they had with merck's structure suggested they could put at substituent at the bottom of pyrrole (R_2 & R_3)
    - did IC_{50} with H, Cl, Br and got smaller numbers than existing drug mevastatin
    - then tested more complex substituents and got IC_{50}=0.025 for Ph and CONHPh groups
      - tested racemic mixtures (faster)
    - to get better understanding, tested the single enantiomer versions and found that the molecule with stereochemistry matching natural compound had much lower IC_{50}
  - pfizer decided to develop the single enantiomer rather than racemic mix to maximize the potency
    - they wanted to use the lowest dose possible bc doctors scared to prescribe statins
    - lipitor would also be the fourth statin drug reaching the market (usually fourth doesn't make much $$)
      - pfizer thought they would make more money if the dose was lower
      - doctors would think its safer if less of a dose is needed
      - lipitor's dose was lower than all other statin drugs
  - lipitor became best selling drug in history; very safe and effective
    - patent expired in 2011, generics now available (price didn't go down much tho)
- assembly to make lipitor:





living catalytic factory genetically engineered e coli. with 2 enzymes from diff species





done to get around the patent

TOPIC 7: ULCERS

November 19, 2020 6:12 PM

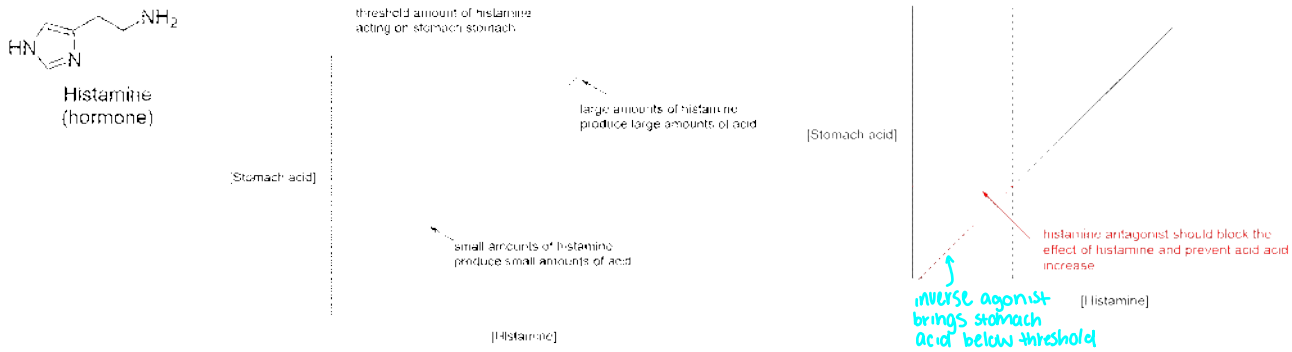
ULCER DRUGS

- before 1960, ulcers where dangerous and often fatal; sig impact on quality of life
- in the 60s, treated ulcers with antacids and diet changes (reduce amount of stomach acid)
- 1976 H2 antagonists (drugs) were introduced
  - lower acid prod in stomach by controlling the regulatory system that regulated acid prod
- 1989; proton pump inhibitors; drugs that inhibit enzyme that makes acid in stomach
  - prevent acid prod in the stomach
- 1984; discovered bacteria that caused most ulcers (H. pylori); leads to a cure for the condition
- ulcers are often fatal; perforations in the stomach often bleed and the blood will pass through digestive system
  - because the bleeding is internal, no blood visible and patient doesn't know they are bleeding to death
- prescription stomach meds are a better way to treat stomach acids than antacids
- theory of ulcer formation in the 60s:
  - stomach contains acid, acid is corrosive
  - therefore excess acid causes corrosion of the stomach liner (ulcer)
  - so the research of the 60s was based on reducing stomach acid
- antacids treat only the symptoms; not for long term use (become problematic)
  - doesn't treat the cause of excess acid, just get rid of excess acid; they are just bases that neutralize the acid
  - examples:
    - sodium bicarbonate and calcium carbonates (both rocks) used in antacids
      - these antacids are basically flavoured chalk
    - $Mg(OH)_2$ (a rock) is in milk of magnesia, an antacid that is a laxative (side effect)
    - metal hydroxides, $Al(OH)_3$ has a side effect of constipation
    - maalox (an antacid) has both $Mg(OH)_2$ and $Al(OH)_3$ so laxative + constipation = perfect??
  - pepto bismol (the pink drink) is not an antacid, its acidic
    - an antibacterial poison, treats diarrhea
  - the metal salts in antacids cause problems
    - body doesn't metabolize metals well so they often build up in body (kidneys etc)
    - so only works short term
- a safer way to treat ulcers is to reduce acid production
  - can use organic compounds (safer than metals long term)
  - address underlying cause directly
- market potential analysis in the 60s:
  - chronic condition; drugs treat not cure
  - serious condition; fatal and bad quality of life
  - large # of patients & developed world, patients can pay

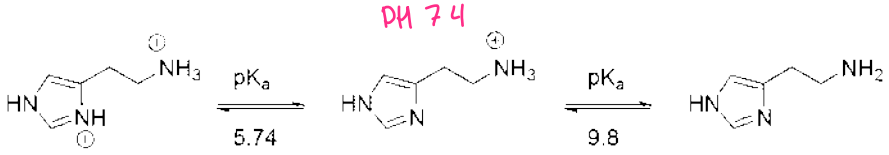
H2 ANTAGONISTS

- cimetidine aka tagamet
  - project started in 1964
  - early example of rational drug design
  - based on animal experiments
    - administering histamine to animals resulted in increased stomach acid prod
      - inc [histamine]=inc [HCl]
    - suggested that histamine (hormone) helped to regulate acid production
    - a selective antihistamine would result in less acid being produced
- there was no structural info available so the histamine receptor wasn't known (high risk project); so needed proof of principle
- histamine antagonists aka antihistamines were possible bc used as allergy meds since the 30s

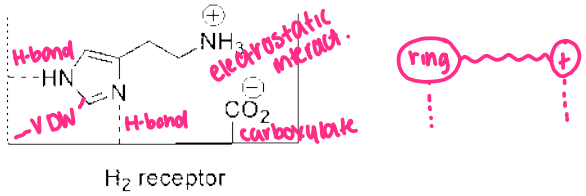
- if histamine is a messenger hormone, there is a receptor; can make an **agonist or antagonist**



- histamine still produced, but gives a weaker effect bc of antagonist (less acid production)
- body always has a small amount of histamine (threshold amount), always a little stomach acid
  - an **antagonist** will keep the level of stomach acid at the threshold amount
  - an **inverse agonist** will reduce stomach acid production below threshold
- high risk project; no evidence that histamine antagonist would work
  - didn't know if histamine was even a hormone, maybe it acts on another system that affects acid prod
  - theory was based on the animal experiments that showed histamine inc acid prod
  - no proof that a baseline level of histamine existed, educated guess
  - antihistamines did exist for allergy, nausea etc but they didn't affect stomach acid
    - ex: ingredient in benadryl is a histamine antagonist but doesn't affect stomach acid prod
    - so caused concern if it would work or not
- project began with histamine agonists and turn it into antagonist
  - change conformation of receptor by changing binding shape
    - agonist binding: normal receptor shape change = biological response
    - antagonist binding: abnormal receptor shape change = no biological response
  - histamine agonist in blood vessels was available
    - H1 receptor; agonist causes vasodilation and lowers BP
  - theory: histamine agonist in stomach
    - H2 receptor; agonist inc HCl production
  - because agonists and antagonists bind in the same place on the receptor, can alter the structure of one via a series of experiments to get the desired effect
    - works best when u know structure of receptor
    - without structure, build a model of binding (pharmacophores)
- to measure agonist behaviour:
  - measure [HCl]
  - administer drug
  - measure [HCl]
- to measure antagonist behaviour:
  - measure [HCl]*
  - administer drug
  - measure [HCl]*
  - administer histamine
  - measure [HCl]*
  - * these numbers should be similar
- rational drug design using histamine structure
  - has 2 protonation sites
  - at pH 7.4, histamine is charged (+)

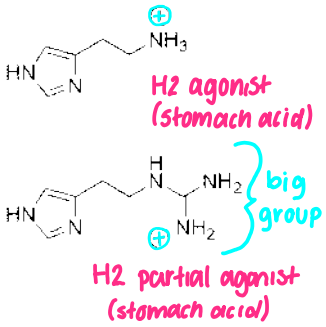


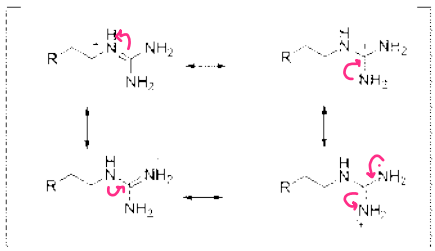
- initial histamine agonist binding hypothesis:

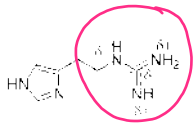


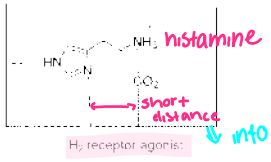
H_2 receptor

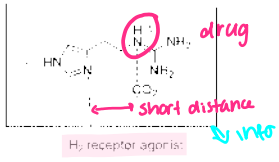
- found that changing the amine changes agonist pattern
  - changing right amine to a big group makes it a partial agonist
  - recall: two types of partial agonists
    - agonist binds to receptor and prod non-ideal shape change = weak signal
    - agonist capable of binding to receptor in more than one way
      - one binding mode give **agonism**, the other gives **antagonism**
      - histamine drug was this type of partial agonist
  - explanation for the 2 binding modes is charge spread
    - due to resonance, positive charge is shared by all four atoms of the group

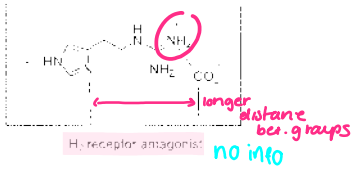






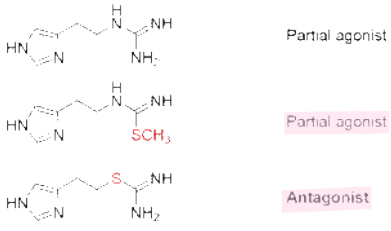


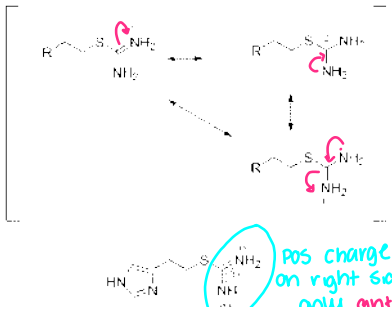


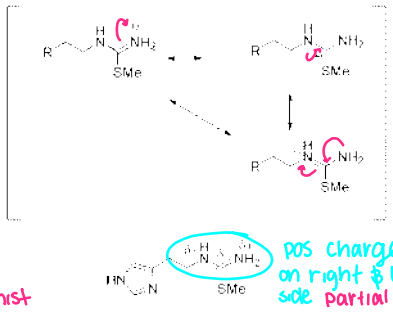


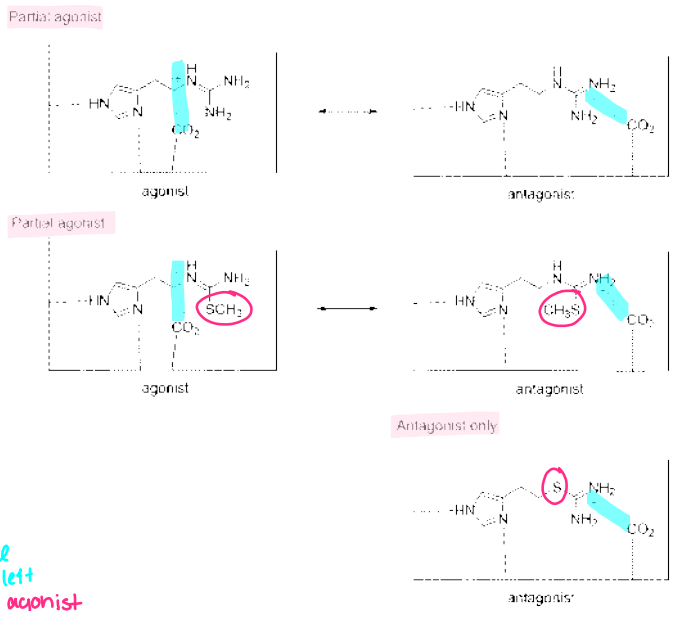
receptor changes its shape for the electrostatic interaction

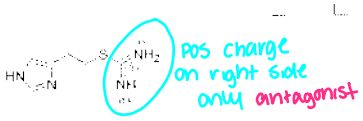
- increasing the distance bet pharmacophores switches an agonist into an antagonist
  - drug forces receptor into diff shape; not able to transmit information
- did amine SAR to increase charge spread and make a complete antagonist (vs just a partial agonist)
  - sulfur doesn't participate a lot in resonance bc large atom and doesn't form pi bonds easily
  - adding sulfur limits locations of pos charge; sulfur doesn't take on positive charge
    - don't often see double bond form bet C&S (C=S)
- longer chain (greater distance bet aromatic ring and positive charge) gives pure antagonism

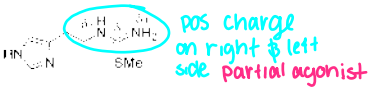


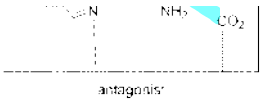




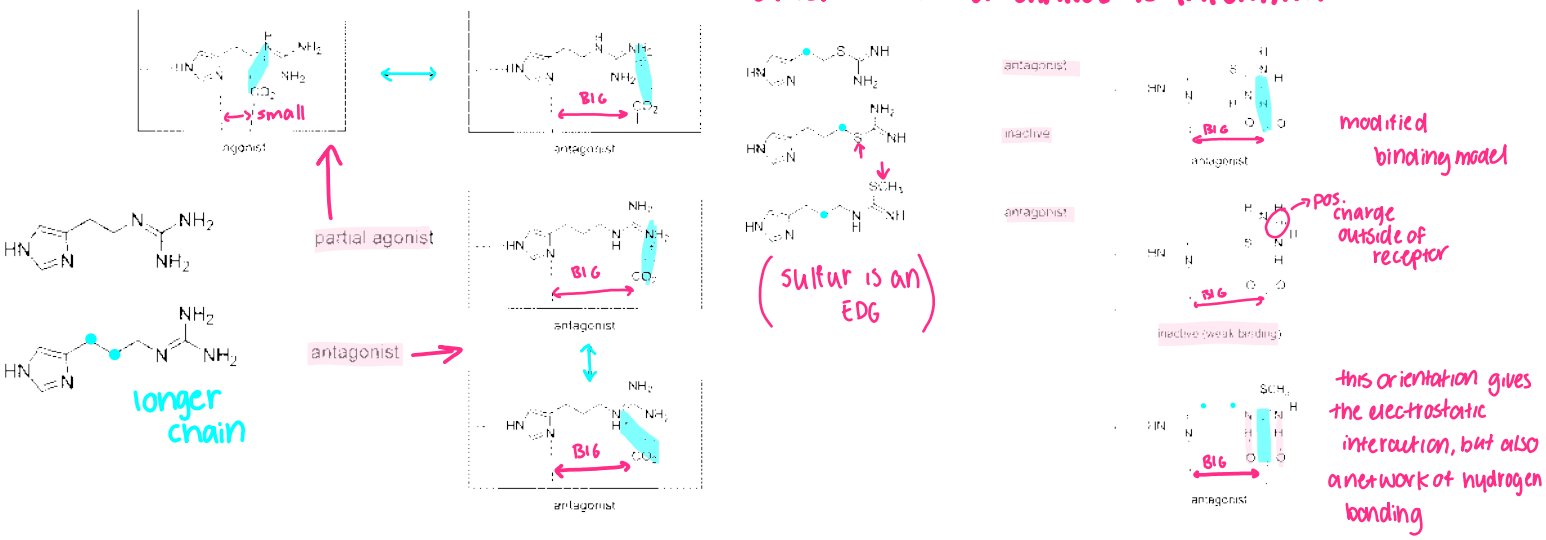



pos charge
on right side
only antagonist

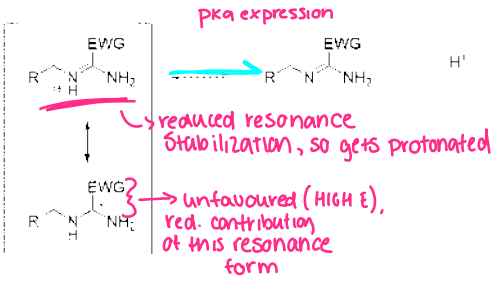

pos charge
on right & left
side partial agonist

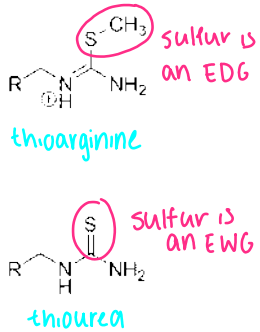

antagonist

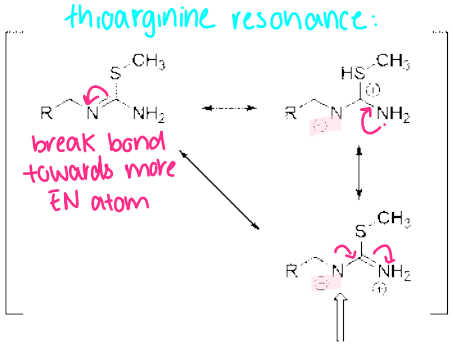
ORIENTATION OF CHARGE IS IMPORTANT.



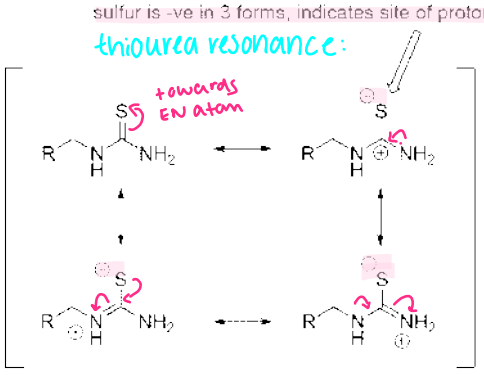
- while electrostatic interaction depends on distance, hydrogen bonding is only directional
- opportunity for a neutral molecule rather than a pos charged molecule
  - add an EWG to prevent formation of + charge (control protonation) and facilitate h-bonding
- use sulfur to control protonation; can act as a EWG or EDG dep on hybridization
  - can change charge without changing sterics
  - S=C makes sulfur an EWG
- to determine whether a molecule is charged at pH 7.4:
  - first identify protonation site
    - use resonance forms to det where there is a neg charge, that is where protonation occurs
  - use pKa expression and look up in pKa value in a table
    - molecule is protonated when pH < pKa (if pKa greater than 7.4)
- thiourea (molecule with C=S EWG) is neutral at pH 7.4
  - can only hydrogen bond, no electrostatic interaction
  - works as an antagonist

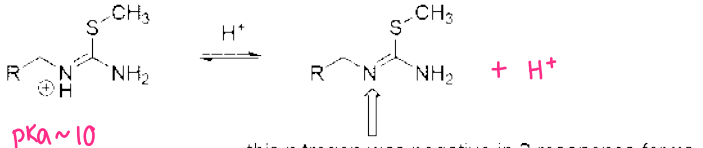




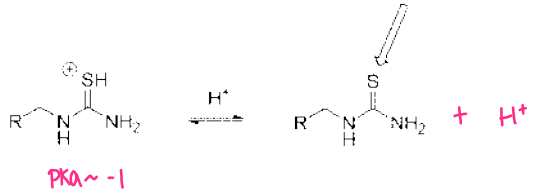


this nitrogen is -ve in 2 forms. indicates site of protonation is on nitrogen

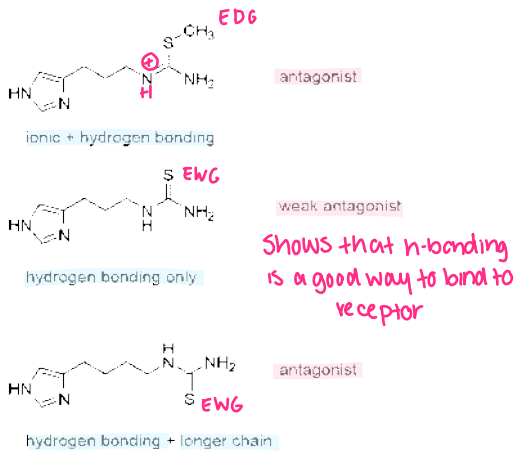


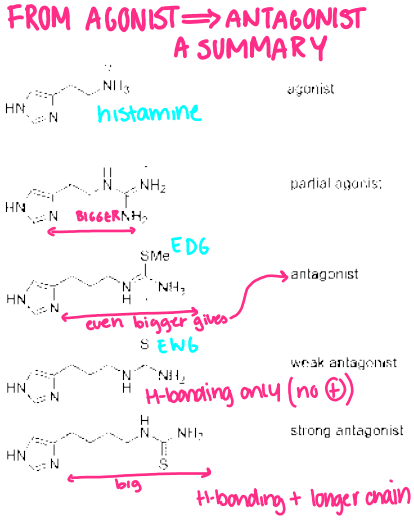


sulfur was negative in 3 resonance forms - site of protonation

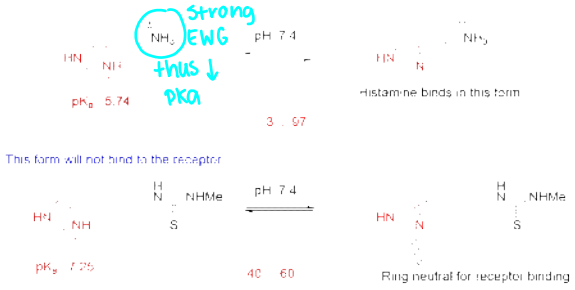


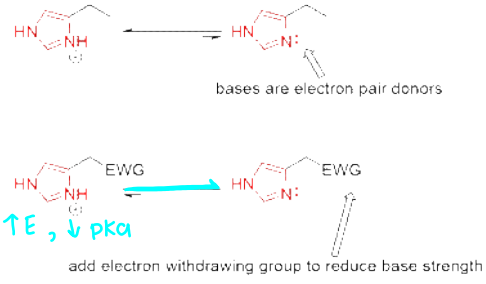
- groups that can only hydrogen bond (no electrostatic interaction bc neutral) work as antagonists:



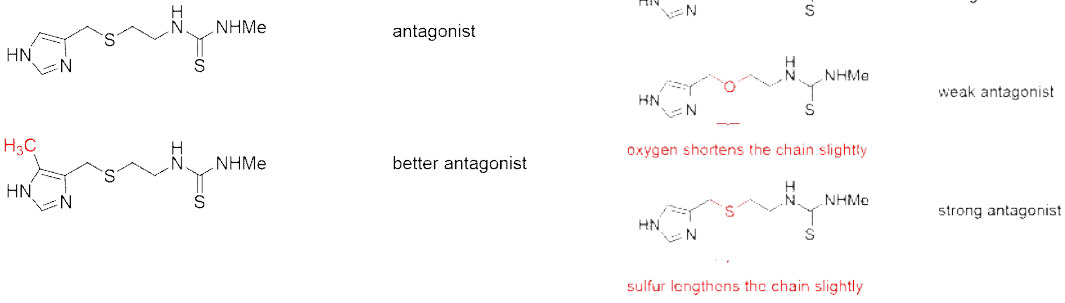


- histamine structure suggests ring must be neutral
  - at pH 7.4, 97% of histamine molecules carry only one positive charge (monocharged form)
  - neutral ring form binds to receptor (VDW interactions)
- the activity of the drug created was not great
  - the pKa of the ring on drug was higher than histamine ring's pKa (even though its the same ring)
  - this made the pKa of the ring close to pH and only ~50% of drug is in monocharged form (neutral ring and pos charge on right that allows receptor binding)
  - to increase activity of drug, need to force more of the drug molecules into neutral ring form; need to make ring less basic
- bases are electron pair donors (get protonated and then have pos charge)
  - can attach EWG to the ring
    - reduces ability of nitrogen to donate electrons (makes it a weaker base) and shifts ring to the neutral form
    - the EWG also destabilizes the positively charged form (makes (+) even bigger, thus high E and unstable molecule)

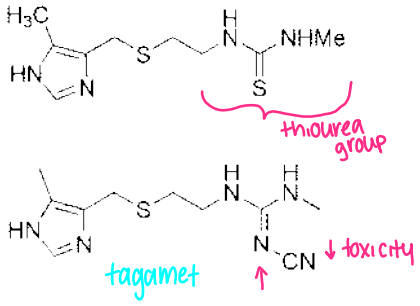




- EWG CH_2 isostere makes ring neutral
  - found that adding oxygen shortens the chain slightly (bc C-O bond is shorter than C-C bond); gives a weak antagonist
  - adding sulfur lengthens the chain slightly (bc C-S bond is longer than C-C bond); gives a strong antagonist
- adding methyl group on ring increases selectivity for histamine receptors in stomach vs histamine receptors in other parts of the body (& controls rotamers?)

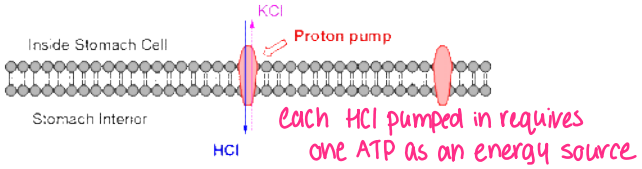


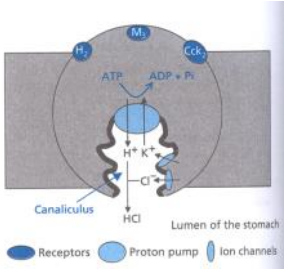
- toxicity problems in clinical trials:
  - kidney damage
  - reduction in white blood cell counts
  - thioureas known to cause toxicity issues in other drugs
    - to fix this, replace thiourea group with bio-isosteres of sulfur
    - NCN bioisostere gave the best drug activity (tagamet)
- tagamet introduced in 1976; first blockbuster drug; $1 billion in sales
  - '64 SAR started
  - '72 cimetidine synthesized
  - '76 tagamet marketed
  - '86 1B annual sales; first drug to do this
  - '94 patent expired; sharp decline in sales bc generic versions available
- other drugs in the same class became available later, most as OTC (zantac, pepcid)



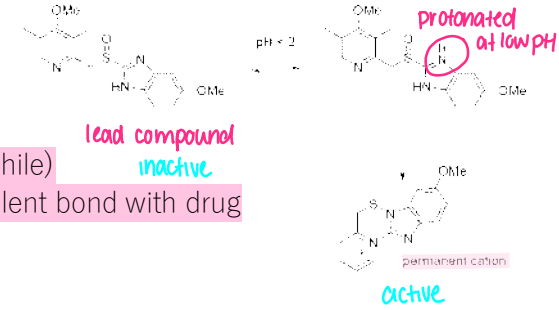
PROTON-PUMP INHIBITORS

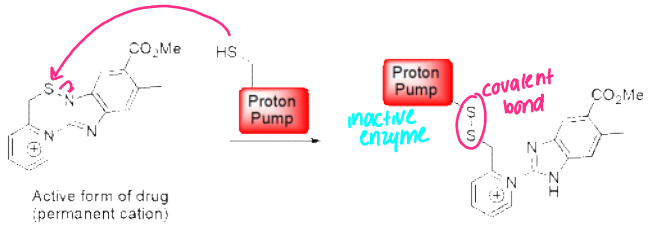
- H2 receptors regulates stomach acid prod by switching stomach acid producing enzyme on/off
- stomach acid can be reduced at the source; an enzyme called the proton pump
- the proton pump is a pore structure that spans the membranes of parietal stomach cells
  - moves protons (HCl) from inside the cell to the stomach interior
  - moves potassium (KCl) from the stomach into the cell to balance this



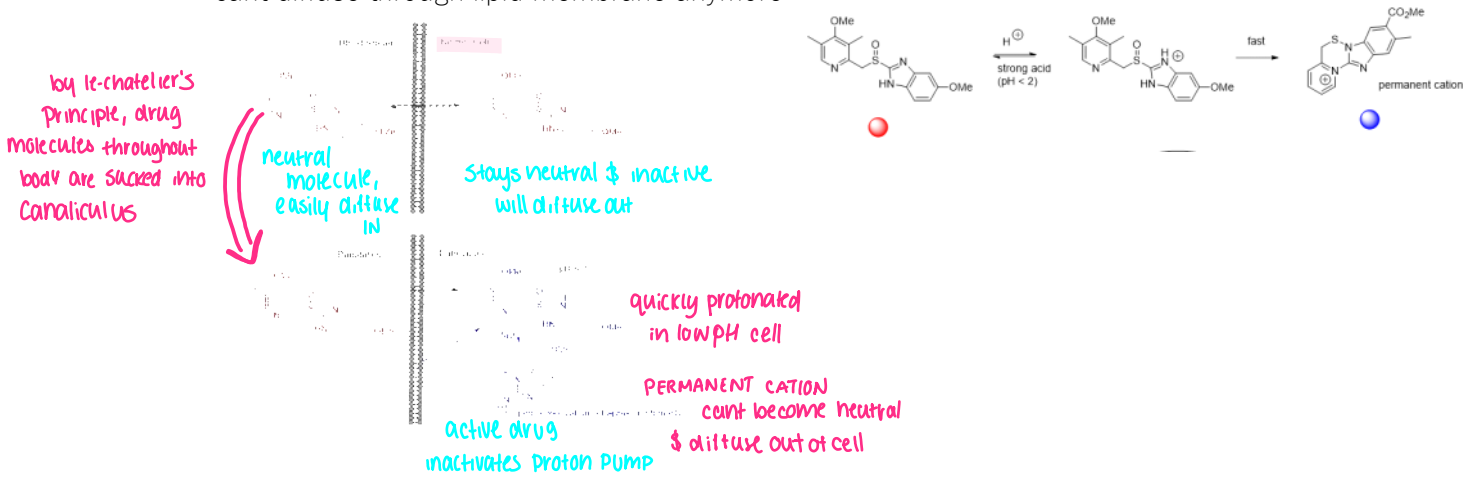


- a proton pump inhibitor is in theory more effective than H2 antagonists
  - proton pump inhibitor would stop all acid production
    - shutting down the machine instead of a control system
  - H2 antagonism only gives partial reduction in HCl
    - H2 antagonist blocks histamine
    - but gastrin (hormone) also stimulates increase in stomach acid
    - and Ach (neurotransmitter) acts like a hormone to increase stomach acid
- proton pump enzyme is an ATPase; converts ATP to ADP
  - hydrolysis of ATP provides energy to move H^+ from cells to stomach lumen and K^+ from stomach to cells
  - ratio of H^+ in cells to H^+ in stomach is $1 : 10^6$
  - ATP provides energy to drive against H^+ gradient
  - inhibit ATP hydrolysis to inhibit the enzyme
- proton pump enzyme discovered in 1977
  - found an inhibition mechanism (inhibit ATPase part of the proton pump)
  - used HTS to identify lead compound
- inhibitor is in an inactive form, gets activated at pH<2
  - gets protonated and makes a permanent cation; drug now active
- inhibition mechanism:
  - proton pump has cysteine (SH) side chain in active site (nucleophile)
  - the cysteine opens one of the rings of the drug and forms a covalent bond with drug
  - this makes the proton pump inactive

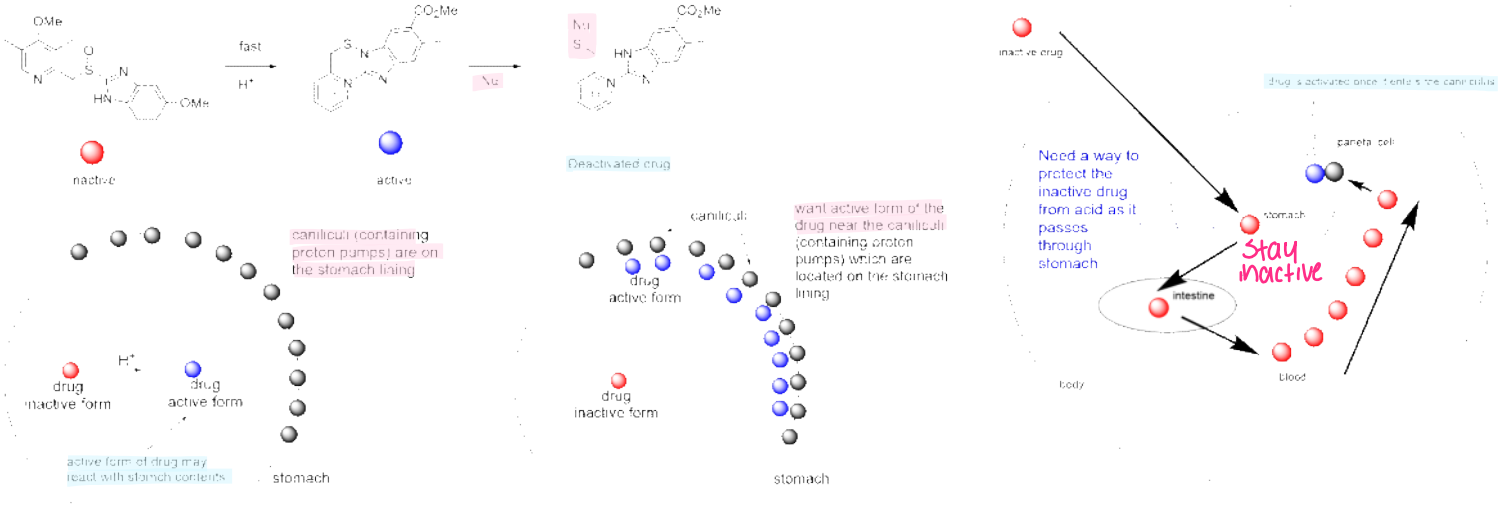




- bioavailability gives selective amplification; so the drug is very clean, low toxicity
  - basicity of benzimidazole is low; only protonated below pH 2
    - the only place in the body this acidic is in the canaliculi
    - canaliculi=pores in stomach parietal cells than contain the proton pump
  - inactive drug circulates in body but can't be protonated
  - protonation in canaliculi traps activated drug in canaliculi
    - cant diffuse through lipid membrane anymore



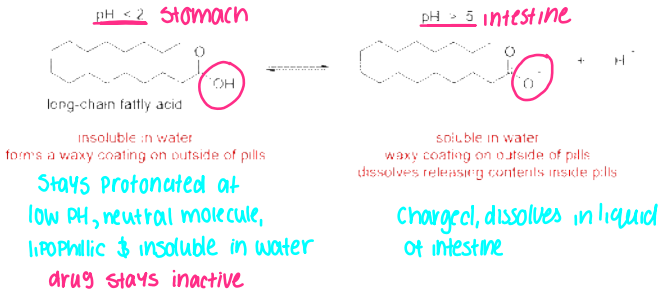
- need to design a pill for selective delivery to the blood while avoiding stomach acid on its way to the blood
  - stomach acid converts the drug to a very reactive compound (active permanent cation, a v good electrophile)
    - will react with any nearby molecule that is nucleophilic
    - reaction permanently "disables" the drug
  - if drug is activated in stomach, it still needs to diffuse to the stomach lining where the canaliculi is
  - but if any food in stomach has a nucleophile, it will disable the drug before it can get to canaliculi and inhibit proton pump
  - we want drug to be transported near stomach lining and then be activated (instead of being activated in the middle of the stomach)
  - if the drug is in the blood (inactive form) it would get secreted into stomach at the stomach lining and then get activated by the stomach acid at the lining where parietal cells/canaliculi are
  - must avoid acid activation in stomach lumen, need to protect pill from stomach acid as it goes to blood



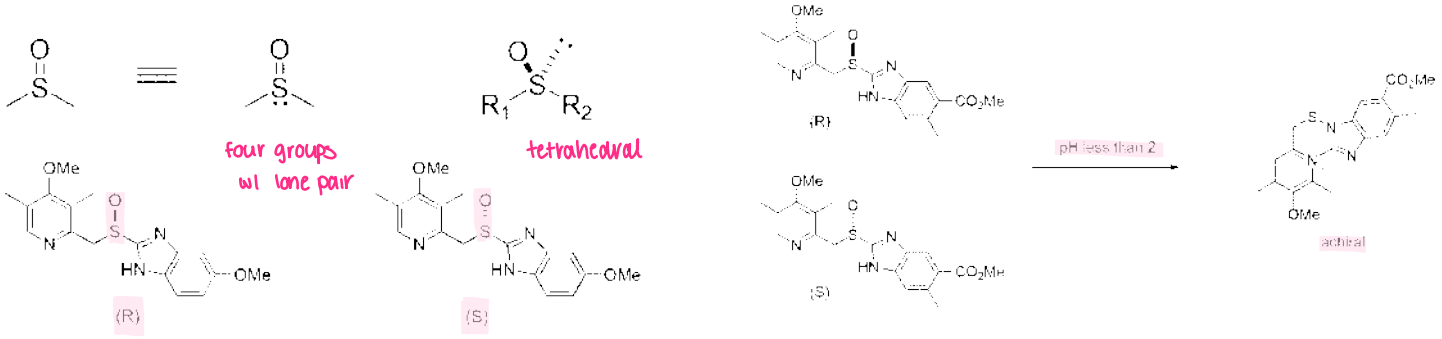
- to make drug stay inactive as it passes stomach, pills have an enteric coating
  - specialized coating stable in acid; prevents pills from dissolving in stomach
  - coating dissolves at higher pH in intestine; drug then dissolves and absorbs in intestine

- some things used include fatty acids, waxes, plastics, shellacs

- fatty acid enteric coating mechanism:



- this drug=omeprazole (aka prilosec)
  - launched in '88, $6.2 B in sales in 2000
  - patent expired in 2001, wanted to protect their market share (bc they were making billions)
    - generic companies would be allowed to make and sell the drug
  - they decided to protect their market share by using a chiral switch
- single enantiomer drugs are hard to make
  - cheaper and easier to make chiral drugs as racemic mixtures; most chiral drugs are made and sold this way
  - each enantiomer acts differently in the body (bc our bodies are chiral environments); efficacy, side effects, metabolism
- chiral switch to protect their market share:
  - patent and sell drug as a racemic mixture (omeprazole)
    - requires that both enantiomers are safe
    - only one enantiomer needs to be active
  - just before patent expires, patent single enantiomer form; often patent is based on manufacturing process
  - FDA considers single enantiomer as a new formulation, don't have to go through all the other testing
    - considering a new form of the drug
  - has only been successful a few times
    - most chiral drugs today are developed as single enantiomers, more technology today
- need to show an advantage, new version is likely more $$ than existing drug
  - why should consumer pay more for single enantiomer version
  - need to convince doctors to switch to more expensive drug
  - need to show advantage for the FDA
    - but FDA considers this a new formulation, doctors are the big problem
- the drug has a sulfoxide, which is stereo genic and can exist in different conformations
- but acid converts both enantiomers into the same molecule
  - R & S isomers have the same biological activity
    - but have small differences due to metabolism
  - get converted into the same achiral molecule at pH<2
- the enantiomers are metabolized at different rates
  - the R enantiomer is metabolized slightly faster than the S isomer
    - liver converts some of the R form into the S form
  - the S isomer lasts a little longer in the body
    - in principle, requires less frequent dosing in the single S enantiomer drug than racemic drug



- esomeprazole (aka nexium) (approved 2001) contains S isomer only
  - advertised to doctors to promote the "improved" drug
  - showed that new drug lasts longer in body; lasts 15 hrs vs omeprazole (racemic version) lasts 12hrs
    - kinda false advertising tho bc they actually administered the drug in a higher conc than the old one
    - also the patient would still have to take 1-2 pills per day with both drugs (24 hrs in a day)
  - benefit is actually to the company, not the patient
  - charged a higher price for the new drug (more than double the racemic version and almost 10x the generic)
- in principle should require less frequent dosing but the numbers don't reflect this
  - S isomer lasts a little longer in the body, but how much longer?
  - dosing didn't actually change

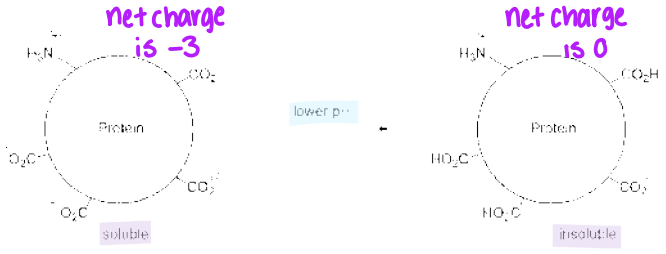
TOPIC 8: INSULIN

November 20, 2020 2:25 PM

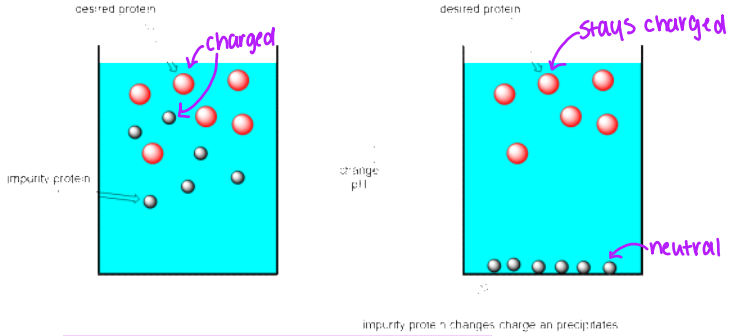
INTRO AND EARLY PURIFICATION

- biologic drugs: vaccines, proteins, antibodies, nucleic acids
  - primarily diff from small molecule drugs by MW (>2000)
  - manufactured using living things (not possible via chemical synthesis)
    - plants, animals, humans, single cells (bacteria, yeast)
  - generally injectable; not able to survive digestive tract
    - other routes (nasal spray, suppository) are possible
  - reserved for diseases that are a quality of life issue (or cancer, untreatable disease)
- increase over the years of biologic drugs bc better technology now to make them
- diabetes; a family of metabolic disorders
  - characterized by high/low blood sugar concentrations; swing in concentrations=big problem
  - one of the leading causes of death in the world
- insulin is a hormone prod by pancreas that regulates blood glucose concentration
  - its presence or absence regulates metabolism of carbohydrates, fats, protein
  - when present, promotes absorption of blood glucose by fat and liver cells
    - cells convert glucose into glycogen (glucose storage polymer) or fat
  - high insulin = low blood sugar = hypoglycemia
    - too much glucose going into cells
    - glucose converted to storage forms: glycogen and fats
    - if insulin too high, glucose unavailable to body cells
  - low insulin = high blood sugar = hyperglycemia
    - glucose can't get into cells
    - storage glucose converted to glucose
    - body also begins to break down proteins (catabolism), can lead to damage
- general types of diabetes
  - type I (10%)
    - pancreas no longer produces insulin
    - starts in late childhood
    - autoimmune component (immune system involved)
  - type II (90%)
    - insulin resistance; pancreas produces insulin but cells don't respond to it
    - autoimmune component
    - linked to obesity
- before 1923, prognosis was death in early teenage years, possible to slightly extend lifespan with diets
- early 20s, banting and best did experiments in toronto
  - surgically altered dogs to enable insulin isolation
    - tied off pancreatic duct certain cells died, leaving behind islets of langerhans (cells that prod insulin)
    - isolated pancreatic extract (insulin) from islets
  - used dog pancreatic extract to treat diabetic dogs (created by removing their pancreas)
- insulin marketed in 1923
  - in 1922 first human experiments failed bc dog insulin was too impure
  - better purification methods gave better results
    - diff animal sources (bovine) gave access to larger amounts of source material
    - bc insulin is found in small amounts in organisms
  - larger amount of insulin makes it possible to perform more purification steps
    - each time you purify, lose a little of the substance
  - early animal choices (by-product of meat prod in slaughterhouses); bovine (cow), porcine (pig), equine (horses), ichthyic (fish)

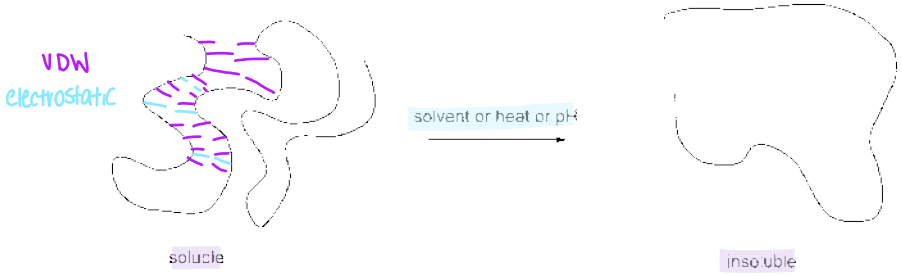
- issues with animal sources: technology back then wasn't good enough to give high purity insulin
  - drugs were a mixture of insulin and other animal proteins
    - easy to remove small molecules (bc v diff in properties), but hard to remove other proteins bc insulin is also a protein and thus v similar properties
    - proinsulin is the protein made by the body to make insulin; similar a.a seq as insulin (immunogenic)
      - proinsulin of different organisms have different structure
      - patients become allergic to a "type" of (pro)insulin
      - need to switch every few years to a different animal source
      - thus it was impo to have access to different animal sources
      - eventually, patients don't have a safe source of insulin
    - glucagon & somatostatin
    - proteases
      - slowly destroy other proteins (insulin) in the drug
      - limits the shelf-life of insulin
- general isolation method in 20s:
  - homogenize pancreases of animal to fracture cells and free insulin from the cells
  - clarification; separate liquids and solids
    - dialyze or centrifuge
    - collect the supernatant (contains insulin)
  - serial precipitation to isolate insulin
    - isoelectric precipitation
    - alcohol denaturation
- clarification of biological suspensions
  - filtration of biological suspensions isnt practical
    - colloids (fine solid suspensions) pass through filters and can clog the filter
    - oily materials clog filters very quickly
  - filtration works in chemical mixtures bc the solid crystals are large (can just use filter paper)
  - dialysis uses continuous diffusion
    - 2 buffers separated by a selective membrane that will allow molecules of a certain size to pass through
      - solvent (doesn't have any of the molecule you want to remove)
      - solution with suspended solids and dissolved molecules
    - molecules dissolved in solution can diffuse through membrane into buffer on the other side
    - since we aren't physically forcing liquid from one side to the other, solid materials remain suspended (don't clog membrane)
    - one method is to put a dialysis bag into the solution and things will diffuse out (teabag analogy)
    - factory method for manufacturing: two pipes with each solution and a membrane between, drug will diffuse to other pipe
      - can continuously refresh the "empty" solvent
  - centrifugation creates artificial gravity and solids sink to the bottom by centripetal force
    - creates a pellet of solids and supernatant (liquid)
- after solids are collected, need to isolate insulin from other components
- serial precipitations to isolate component of interest; serial precipitations in the proper order will purify insulin
  - change chemical environment of the solution
  - this will change properties of some of the molecules to make them insoluble
    - molecules will be precipitated (dissolved in solution to solid)
  - isoelectric precipitation
    - change the pH of the solution using a buffer
    - alters the charge state of proteins
      - charged proteins=soluble
      - neutral proteins=insoluble



        - every protein has an isoelectric point/state (pI): pH where the protein has no net charge (pos charge=neg charge); at all other pH values the protein will have a net charge
        - can det isoelectric pH if you know the pKa's of the side chains on protein
        - can precipitate impurity proteins or desired proteins, depending on the nature of the proteins and the process you're using



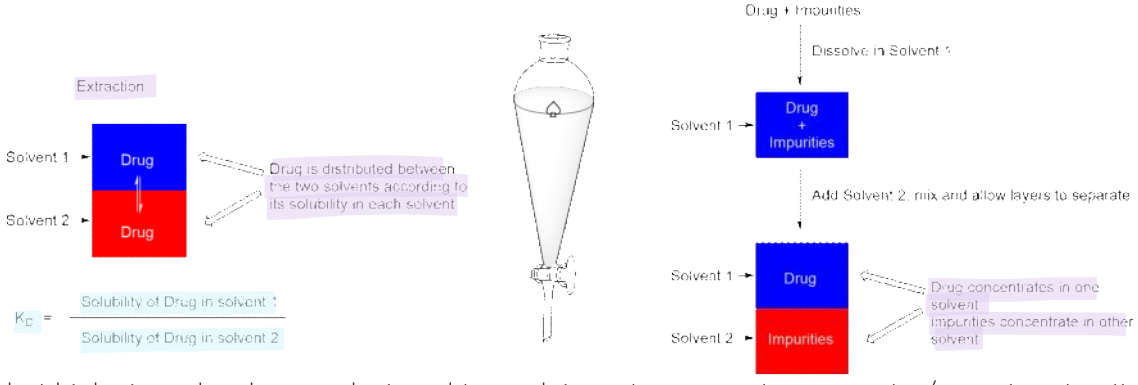
    - protein denaturation; disrupt tertiary structure of protein by changing pH or solvent polarity or heat
        - changing pH; changes charges on certain groups and disrupts interactions holding parts of the protein together
        - changing solvent polarity; disrupts interactions between protein chains
            - non-polar solvents:
                - ex: protein has vanderwaals interaction holding two chains together and an inner electrostatic charge
                - when you add a nonpolar solvent, it can separate the chains held together by VDW, this exposes the inner electrostatic interaction to water molecules which will weaken the interaction and unfold protein
            - using salts to make solution more polar; electrostatic interactions holding protein together will interact more with the solvent than each other and protein unfolds
        - heat will increase kinetic energy; proteins is moving around more which can pull the protein apart in areas and exposes electrostatic interactions to water (used to be protected by VDW interactions)



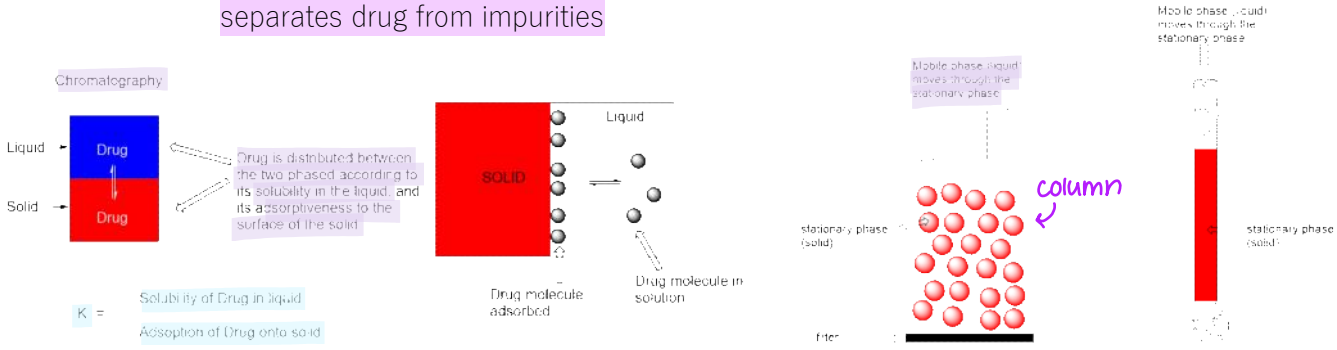
        - unfolding protein changes the physical properties of the protein and makes it insoluble in water so it precipitates
        - usually precipitate impurity protein and keep desired protein in solution bc that way it stay folded up in its shape
        - but can also precipitate desired protein if you have a way to renature it (less common)
- in the early 30s, there was a modification to this purification procedure; came into effect in the late stages of the sequence
- zinc promoted crystallization was discovered in the early 30s
    - insulin forms hexamer crystals in the presence of Zn^{2+}
        - 6 units of insulin would crystallize with zinc ion
    - partly purified insulin is dissolved and crystalized by adding Zn^{2+}; crystals of insulin
    - crystals are then recrystallized, gives much higher purity
        - recrystallization requires that you start with something already mostly pure
    - main impurity after recrystallization is still proinsulin but in lower amounts
        - much lower rates of allergy
        - longer onset for intolerance of species-derived insulin

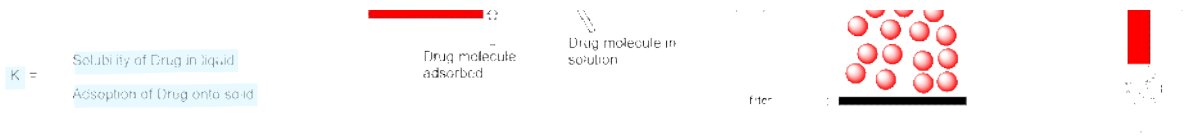
GENERAL CHROMATOGRAPHY

- development of chromatography in the late 50s led to production of higher quality insulin
  - allows a higher grade removal of impurities
  - used in late stages of purification process; requires that most impurities are already gone
  - ion exchange chromatography separates proteins with different ionization levels
  - gel filtration/size exclusion chromatography separates proteins based on size
  - HPLC aka high pressure/performance liquid chromatography methods make the biggest impact
    - change in particle size; use small particles instead of large particles
    - improves resolution
- chromatography is sequential extraction
  - extraction partitions molecules between two solvents (phases)
    - ex: dissolve drug in solvent 1. add solvent 2, drug will achieve an equilibrium between the two solvents
      - K_D=equilibrium constant=distribution constant=solubility of drug in solvent 1/solubility of drug in solvent 2
    - chemicals tends to concentrate in one of the solvents/phases; by separating the solvents you can separate components in the mixture
    - can purify drug based on solubility properties; drug will concentrate in one solvent and impurities will concentrate in the other solvent
    - solvent 1 & 2 are chosen so that they are not miscible with each other
      - usually water and an organic solvent (both phases are liquid); works for small molecules

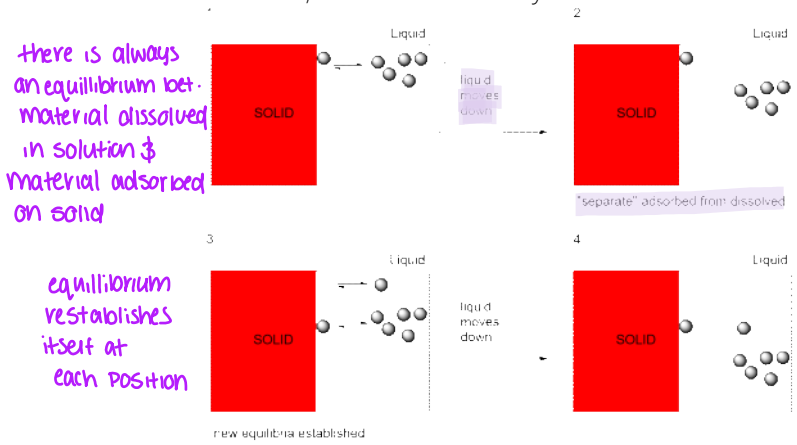


  - but biologic molecules are designed to work in water, so cant use a water/organic extraction
    - organic material will cause biologic material to denature, can't do extraction as purification in biochem
    - use chromatography instead; change one of the phases to a solid
  - chromatography partitions molecules between a solvent and a solid
    - dissolve drug in liquid. expose this to solid and the drug can stick to the surface of the solid
    - K=equilibrium constant=solubility of drug in liquid/adsorption of drug into solid
    - stationary phase=solid & mobile phase=liquid
      - ex: TLC plates where stationary phase is powdered silica gel on the plate and mobile phase is the liquid that moves up the plate
    - to purify on a larger scale, need more than a plate; use a column
    - column is like a pipe, inside it is the stationary phase (solid particles)
    - mobile phase (liquid with drug dissolved) is forced to flow through small particles of solid
    - as the drug (with impurities) passes through the column, there are a series of extractions that separates drug from impurities

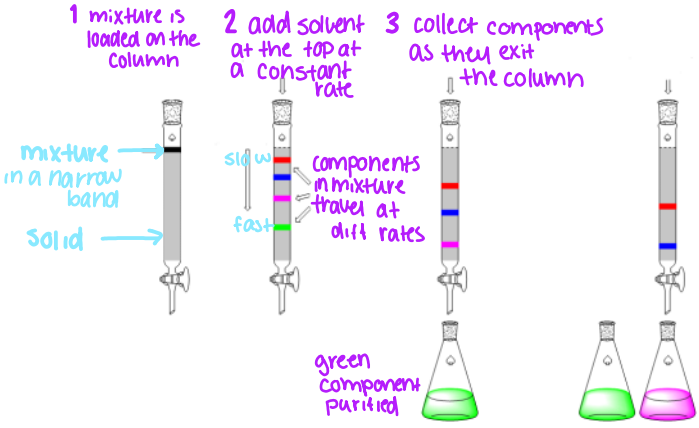



K =
Solubility of Drug in liquid
Adsoption of Drug onto solid
Drug molecule adsorbed
Drug molecule in solution

        - moving solution is in equilibrium:
            - slowly separating the mixture as more liquid is added and moving down
            - if a material is mostly in solution, it will move down the column really fast bc most isnt stuck to the solid
            - if a material is mostly stuck to the solid and only a few are in the liquid, most of it gets left behind on the solid; it will move slowly down the column



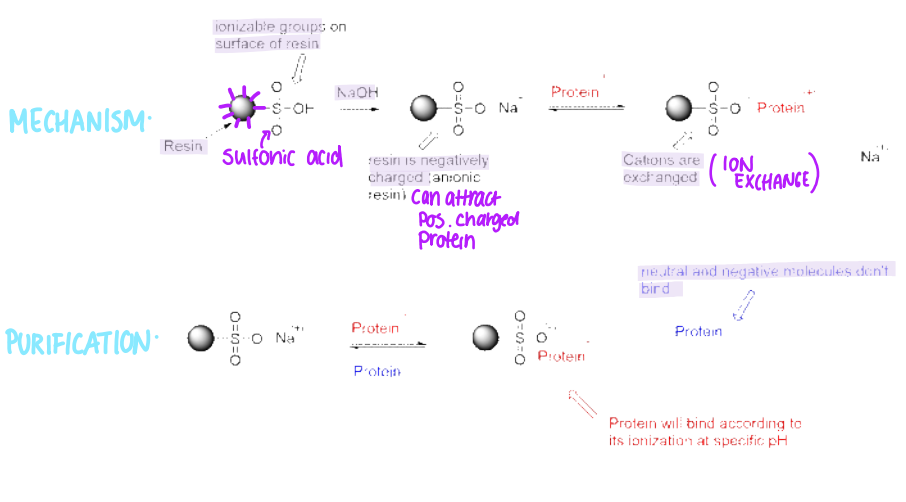
        - column thus functions like a series of repeated extractions
            - molecules continually exchange between solution and adsorption as liquid moves through the column
            - different molecules exchange at different rates; different equilibrium for each molecule type
            - controls how quickly molecules move through the column
                - low adsorption = fast passage
                - high adsorption = slow passage
- column operation:
    - mixture (in solution) is loaded at the top of the column; max concentration and in a narrow zone
        - want everything to start at the same place
    - solvent is added at the top and removed from bottom
        - continuous flow of solvent through column
    - mixture separates as it moves through column
    - collect different molecules as they exit
        - molecules will take diff amount of times to exit column depending on their rate

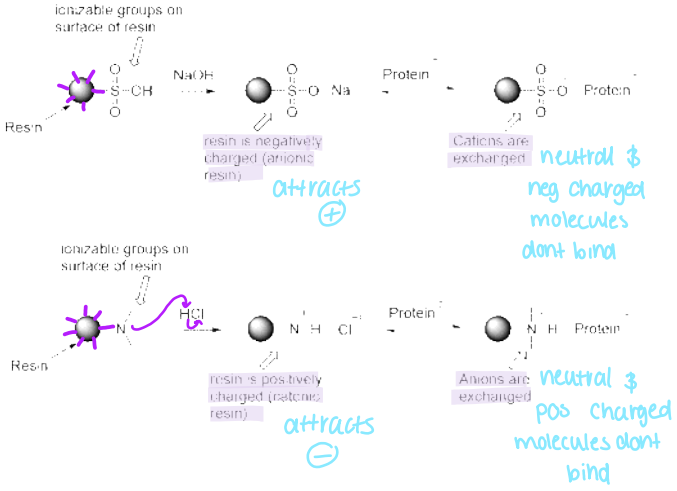


- biological compounds are colourless, cant see them so don't know when components are exiting the column
- in the lab, collect a set of fractions of approximately the same volume
    - leave each flask to collect for a set period of time
    - then test each fraction to find the component you are looking for
- large scale production uses detector at the column outflow
    - changes fractions when detect new molecules exiting
    - don't need to test each fraction, you know which flask has each component

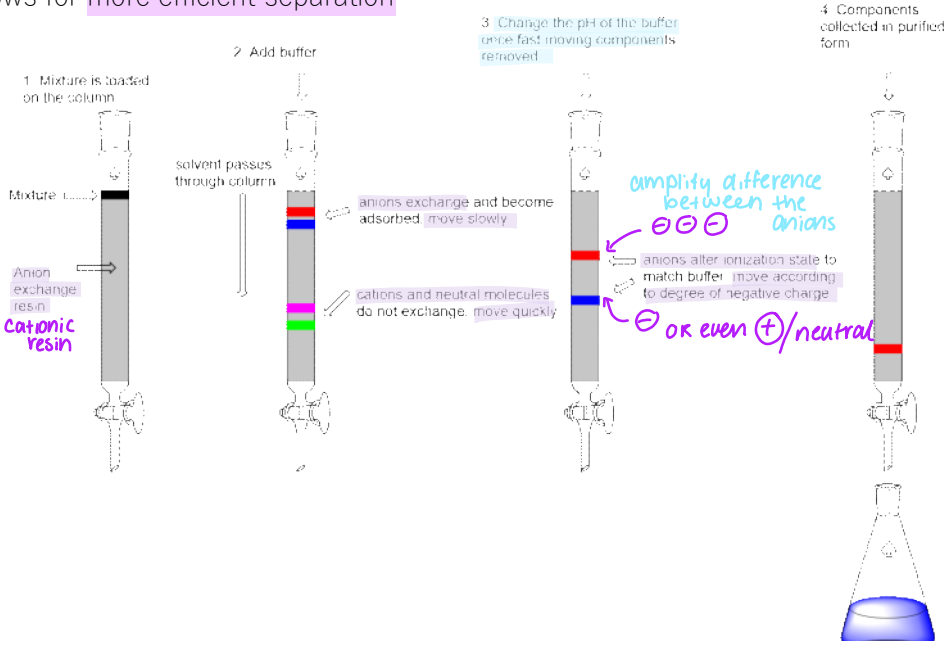
ION-EXCHANGE CHROMATOGRAPHY

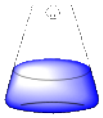
- electrostatic interaction with resin (solid) in column differentiate different types of molecules
  - resins are engineered/manufacturing, basically plastics
  - resin properties can be adapted for use in chromatography
- molecules adhere to resin according to their charge:
  - highly charged stick tightly
  - partly charged stick partly
  - neutral don't stick
- if molecule doesn't bind resin, low adsorption = fast through column
- if molecule binds resin, slower through column
- by controlling pH, can control state of charge of various proteins
- various types of resins can be used
  - anionic resins aka cation exchange resin
    - negatively charged resin that exchanges positively charged molecules
    - ex: sulfonic acid treated with NaOH
  - cationic resins aka anion exchange resin
    - positively charged resin that exchanges negatively charged molecules
    - ex: amino groups treated with HCl





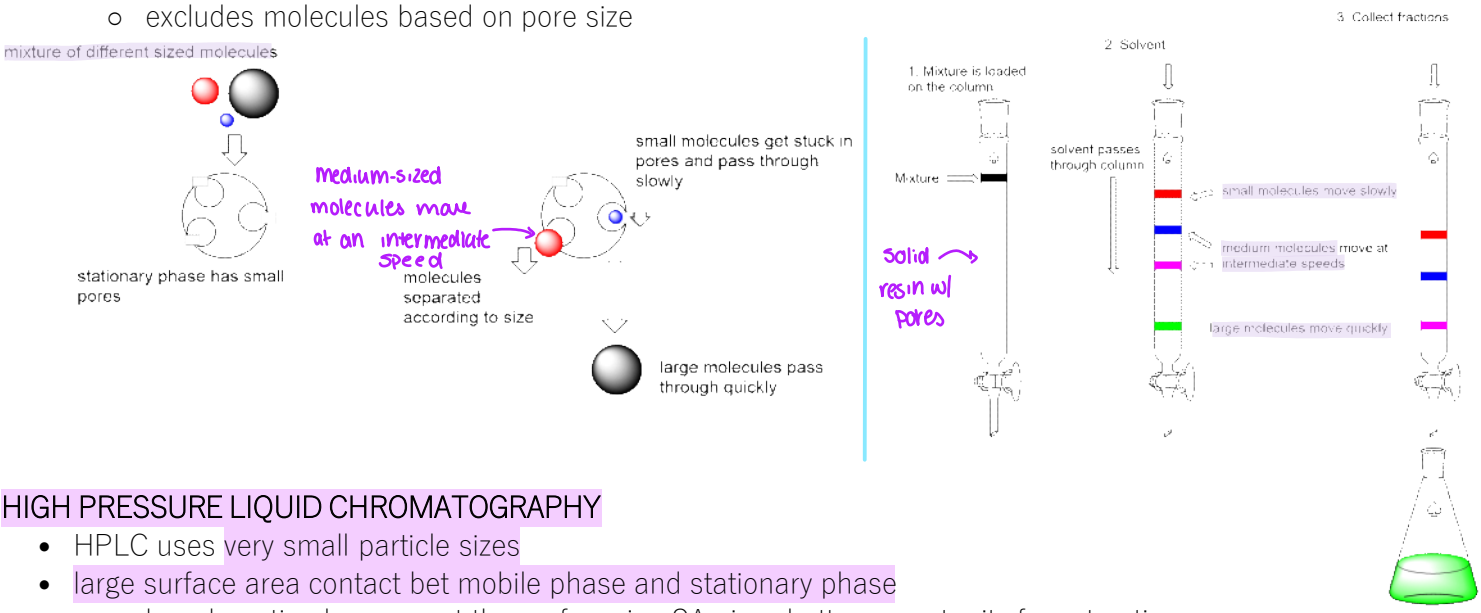
- control charge on proteins by changing the pH, so the mobile phase is usually a buffer
- possible to change pH during the procedure
  - alters ionization as the separation proceeds
  - allows for more efficient separation





SIZE EXCLUSION GELS

- stationary phase (resin) contains fixed-size pores
- small molecules fit inside the pores; move slowly through the gel
- large molecules don't fit into the pores; move quickly through the gel
    - excludes molecules based on pore size



HIGH PRESSURE LIQUID CHROMATOGRAPHY

- HPLC uses very small particle sizes
- large surface area contact bet mobile phase and stationary phase
    - bc adsorption happens at the surface, inc SA gives better opportunity for extractions
    - can increase resolution, better separation with bigger SA for contact
- requires high pressure to force the mobile phase through stationary phase
    - with large particles, there is a lot of space for liquid to move when you push liquid into the column
    - tightly packed small molecules give little space for liquid to move through, need more pressure
- columns have thick steel wall to resist pressure
    - force liquid through bottom, collect at top

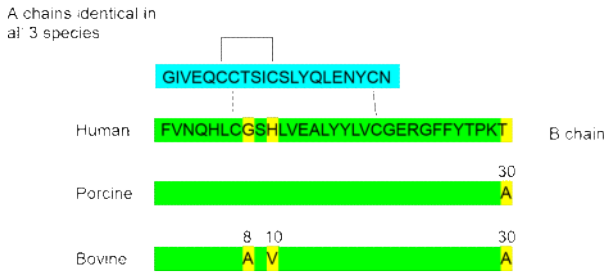
INSULIN PRODUCTION IN THE 70s

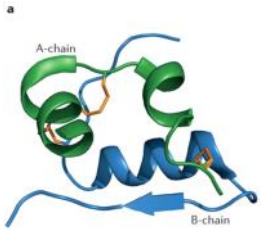
- clarification
- precipitation
- crystallization and recrystallization (zinc)
- HPLC; use either ion exchange or size exclusion
- recrystallization (zinc)
- really pure insulin; side effects due to animal source contamination was rare by the late 70s

MANUFACTURING PROTEINS (INSULIN)

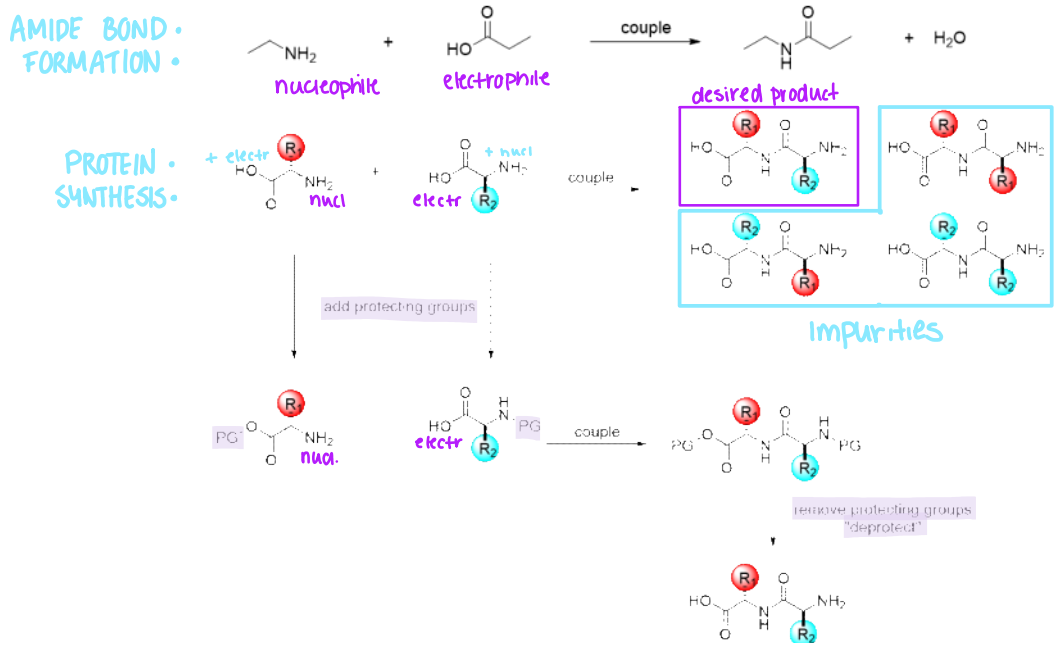
- possible manufacturing methods (source of raw material):
    - extraction from animals
    - chemical synthesis
    - genetically modified organisms
- disadvantages of using animal sources for protein drug manufacture:
    - proteins aren't human proteins similar but not the same (diff a.a. sequence)
    - purity; impurities are biological
        - similar properties, so hard to remove
        - allergy due to other proteins (proinsulin)
        - viruses (infection); small risk tho
    - need large amounts of animals; supply may be limited
        - hormone proteins (like insulin) are in small amounts in organisms
- via chemical synthesis, can make human protein only
    - could solve purity problems bc impurities will generally be small molecules (no protein impurities)
    - easier to separate small/big molecules

- general structure of insulin:
  - small protein composed of 2 chains (A and B) connected by two disulfide bonds (has quaternary structure)
    - A chain is 21 amino acids with an intrachain disulfide bond
    - B chain is 30 amino acids
- human vs animal (porcine and bovine) insulin sequences:
  - A chain is identical
  - animal B chains differ from human:
    - porcine(pig); one a.a difference at position 30
    - bovine (cow); three a.a differences at positions 8, 10, and 30



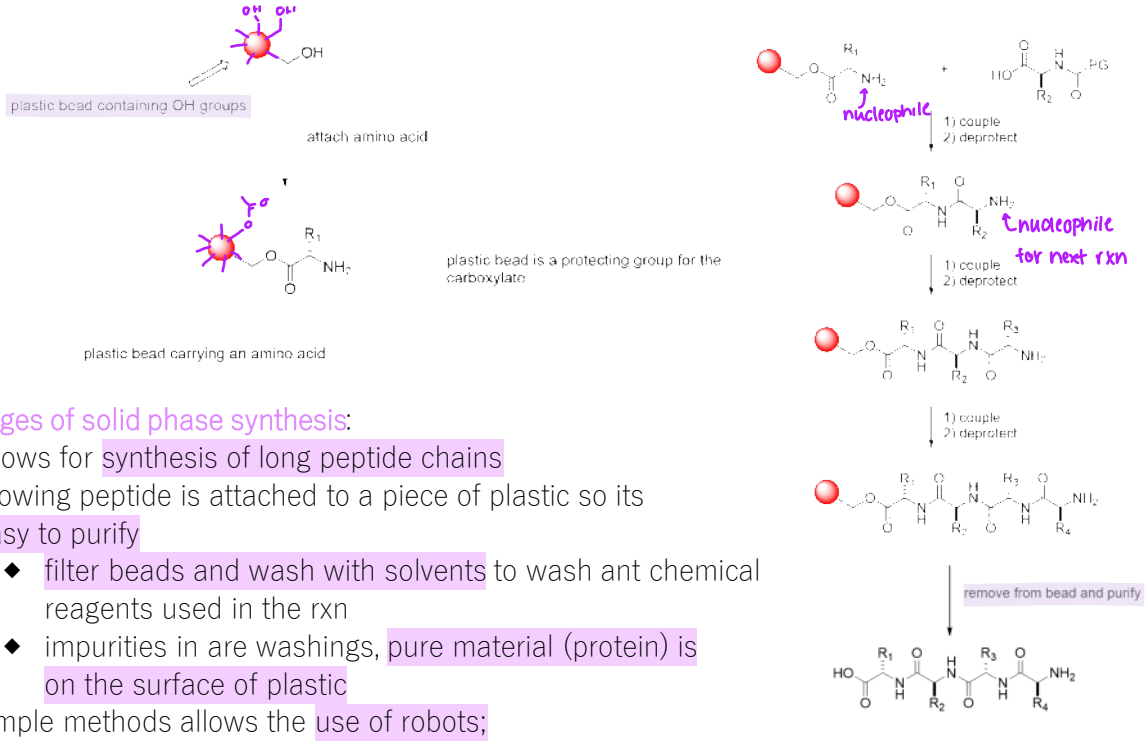


- insulin was the first protein made entirely by chemical synthesis in 1964
- no risk of large protein contamination, can make higher purity insulin
- but chemical synthesis has issues, it works in the lab but wont work to make a drug
- to make protein in the lab via chemical synthesis:
  - mix amine with a carboxylic acid; couple them to make an amide bond
    - works bc just one nucleophile and one electrophile
  - but amino acids have an amino group and acid; when you couple two amino acids they can form four different products
    - two nucleophiles and two electrophiles
    - amino acids react with e/o other in two diff ways or with itself
    - only 1/4 of the product is want you want, the rest is impurity
  - use protecting groups to get selective formation of amide bonds
    - modify each amino acid so it has only one reactive part
    - add protecting group on amino group if you want the amino acid to be the nucleophile and on acid if you want the amino acid to be the nucleophile
    - get only one product (with protecting groups)
    - remove protecting group (deprotection) and you get the desired product




R1
HO
O
N
H
O
NH2
R2

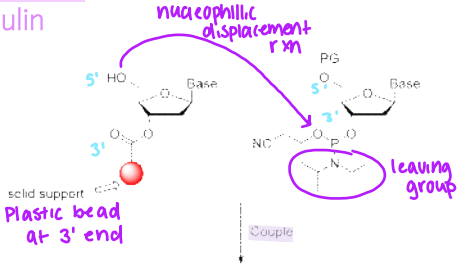
- technical issues as the size of the protein grows
  - as a.a are added and the protein gets bigger, it gets harder to work with
  - hard to get into solution, hard to get it back out
  - so hard to purify when you get above 4 a.a. in the chain
- to overcome this, used technique called solid phase protecting groups
  - protecting group is attached to a plastic bead
    - the plastic carries a functional group (ex: OH nucleophile)
  - one bead/PG binds many chains
  - allows the bead to function in a chemical manner, but bc its plastic, functions as a solid
  - control it easily; can physically pick up and move the molecule

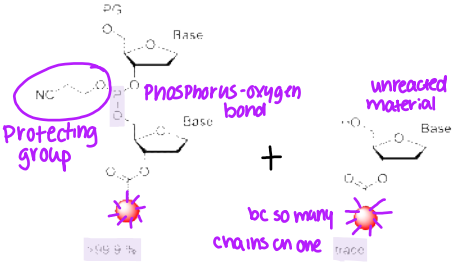


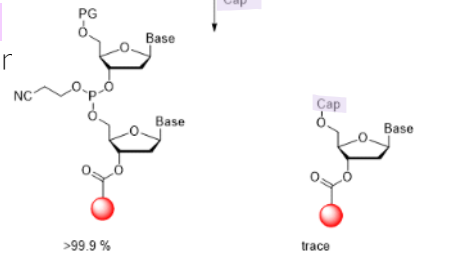
  - advantages of solid phase synthesis:
    - allows for synthesis of long peptide chains
    - growing peptide is attached to a piece of plastic so its easy to purify
      - filter beads and wash with solvents to wash ant chemical reagents used in the rxn
      - impurities in are washings, pure material (protein) is on the surface of plastic
    - simple methods allows the use of robots; automated synthesis of proteins
      - use machine called a peptide synthesizer
      - enter in the computer the sequence of a.a you want to make, and the machine will make it
      - amino acid is made and stuck onto the bead; remove and purify
- this method was used to make HIV protease with all D amino acids (rare in nature)
  - tertiary & quaternary structure was a perfect mirror image to the structure of the natural enzyme
  - was the final proof that tertiary structure is dictated by primary structure only
- method was adapted for automated synthesis of all biological polymers
  - peptides, proteins, DNA, RNA, carbohydrates
  - all methods use solid phase synthesis
- limitations of chemical synthesis of proteins:
  - small quantities only (mg amounts); only used in labs
  - process produces large amounts of chemical waste
    - hundreds of liters of waste per mg of protein made
    - v wasteful to do this on a large production scale; cost prohibitive
  - can only make chains about 20 a.a long
    - longer proteins made by linking chains together via other methods

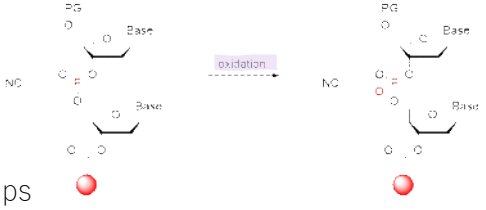
- goals of protein drug manufacture:
  - high purity; limited # of impurities and small amounts (chm synthesis gives good purity)
  - want human proteins; work better and less chance of allergy
  - large amounts of protein; tons of protein (chem synthesis isn't practical)
  - low cost (chem synthesis isn't practical bc waste=cost)

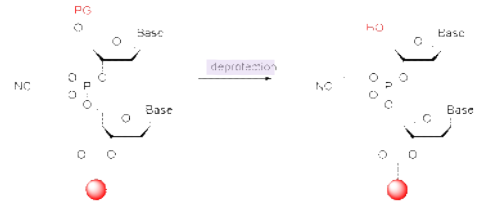
- use a genetically modified organism for protein synthesis; engineer a living factory to make the drug
  - the first synthetic gene was made in 1978:
    - somatostatin, a 14 a.a protein
    - gene (DNA) was made using a machine
    - gene was expressed in E. Coli
      - isolated somatostatin (protein)
      - first recombinant protein ever made
        - made by combining DNA from different sources
  - recombinant insulin was made in 1979
    - used chemical synthesis (machine) to make DNA corresponding to the A & B chains of insulin
    - expressed each gene separately in E. Coli to produce the A & B peptides; easier purification
    - combined the A & B peptides chemically to make recombinant insulin
  - use a DNA synthesizer to chemically make DNA
    - enter sequence of DNA, and it makes it
    - does oligonucleotide synthesis (coupling)
    - in biological systems, nucleic acids are synthesized from 5' → 3'
      - use 3' OH as a nucleophile
    - in chemical synthesis, nucleic acids are synthesized from 3' → 5'
      - use 5' OH as a nucleophile
      - 5' OH is a primary alcohol, primary nucleophiles are better
      - 3' OH is a secondary alcohol, too bulky
    - use solid plastic bead as protecting group on 3' end
    - after coupling, some trace 5' OH nucleophiles remain unreacted
      - need to cover these up bc will generate a nucleic acid impurity that is v hard to get rid of
        - hard to separate multiple DNA strands
        - must make sure only 1 sequence is made
      - coupling rxn is followed by a capping rxn, add a small group (cap) to block the position
    - coupled phosphorus in a lower oxidation state than normally found in nucleic acids (bc coupling reaction is easier to do in low oxidatior
      - must oxidize to proper level, capping is followed by oxidation
    - for next nucleic acid addition, need to remove protecting group on 5' OH
      - oxidation is followed by deprotection to unblock 5' OH for the next coupling rxn
    - reaction sequence is repetitive for addition of nucleic acids:
      1. coupling rxn
      2. capping rxn
      3. oxidization
      4. deprotection
    - after chain is complete, final step is removing all protecting groups
      - remove plastic bead and phosphorus PGs using ammonia

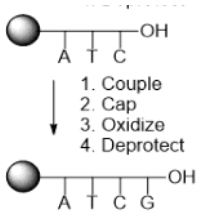


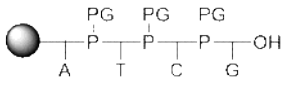


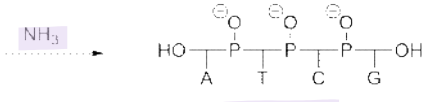


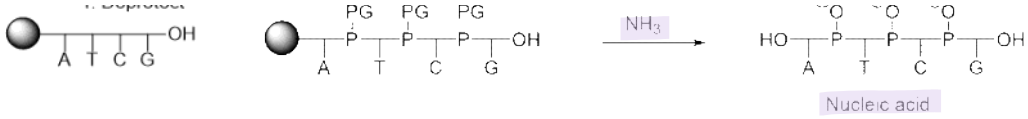








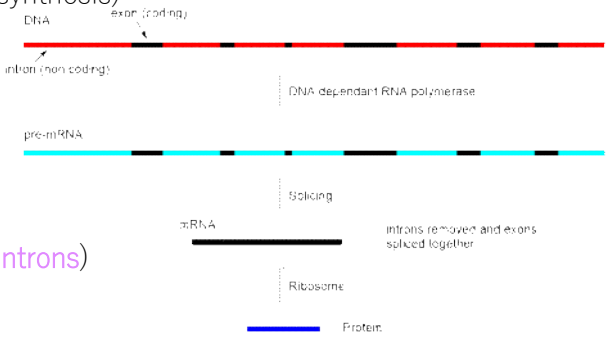



OH
A T C G
PG PG PG
P P P OH
A T C G
NH3
O O O
HO P P P OH
A T C G
Nucleic acid

- the first synthetic life form was made using the DNA synthesizer machine
  - made the entire genome of a bacteria using the machine
  - 1000 gene fragments were synthesized on gene machines
    - each fragment was 1080 base pairs
      - 1000 double stranded base pairs
      - 40 single stranded base pairs on either end of the ds
        - "sticky ends" for ligation
  - entire genome 1M base pairs
  - the gene fragments were assembled using yeast to link fragments into a single circular chromosome
  - the synthetic gene was transplanted to another species
    - removed chromosomes from *M. capricolum* cells; need all other enzymes, protein etc to read and replicate DNA
    - synthetic *M. mycoides* chromosomes implanted into the empty cells
      - LacZ gene was added to synthetic gene as a marker aka reporter gene
        - causes bacteria to turn blue, know which colonies are the good ones
      - blue colonies of *M. mycoides* were grown
    - other than being blue, the synthetic bacteria were indistinguishable from the natural form
- the first artificial chromosome of a eukaryote was made in 2014; chromosome of yeast
  - all DNA sequences were machine made; modified sequence of wild-type for research purposes
    - made synthetic telomeres @ end of chromosome
    - added special sequences to manipulate the chromosome
    - added reporter genes
    - made modifications to specific sequences to study when genes are expressed
  - currently manufacturing all chromosomes of yeast to make the first artificial eukaryote
  - to save labour costs, the work is carried out by undergrad student in a build-a-genome course
  - 6 out of 16 chromosomes are currently completed
- if the method of recombinant insulin synthesis could be scaled up, it could be a better way to manufacture insulin in large quantities using bacteria as a living factory

RECOMBINANT INSULIN

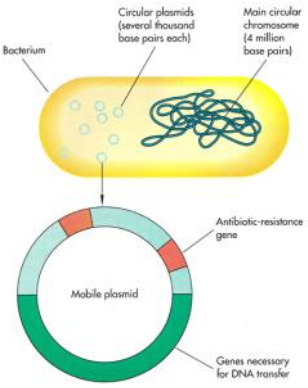
- steps to make recombinant insulin:
  - isolate DNA for the protein you want (mRNA or machine synthesis)
  - insert the DNA into a vector (plasmid or virus)
  - insert the vector into cells (transformation)
  - isolate the cells that express the gene you inserted
  - grow the cells in large quantity
  - isolate and purify the protein you want

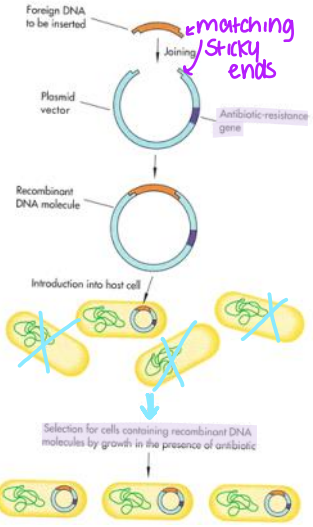


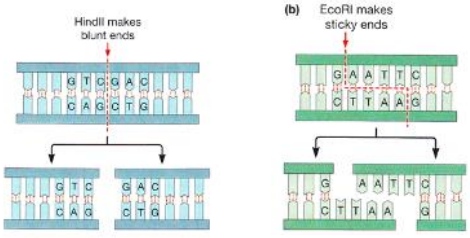
- advantage of chemical synthesis of DNA recombinant work:
  - human genes contain long stretches of noncoding DNA (introns) that are removed during transcription
    - coding regions are exons
  - if you get DNA from source (humans) need to remove the introns to get readable DNA for cloning
  - most humans genes are non-coding DNA
    - ex: factor VIII gene; only 5% of the gene is exons
    - if you put this into bacteria, it would try to read the entire gene
  - bacterial genes don't have introns so mRNA isn't processed to remove introns
    - bc insulin gene was small, the easier way to do this was to make DNA on a machine with no introns
- plasmids commonly used as vectors to make genetic modifications
  - small and easily purified
  - can work with them in vitro
  - some plasmids multiply
    - get multiple copies inside bacteria

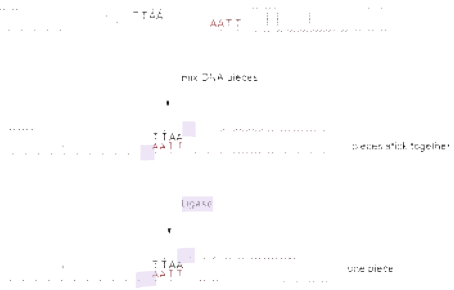
- amplifies amount of protein

- plasmids are autonomously replicating mini chromosomes
  - small, circular DNA molecules found in bacteria
- not connected to main chromosome; some can move on/off the main chromosome, these are called episomes
- episomes can transfer between bacteria; one of the main mechanisms of how resistance/genetic info is transferred
- modern plasmids are fully engineered, commercially available
  - antibiotic resistance gene inserted (used for purification)
    - grow cells in media containing antibiotic
    - only cells that properly express plasmid survive
  - color-producing enzyme for colony identification; reporter gene
  - most have a polylinker for DNA splicing; special DNA sequence containing lots of restriction sites
    - allows you to insert sequences
    - restriction sites are targets for restriction enzymes
    - restriction enzymes cut DNA at specific sequences (restriction sites)
    - can make blunt ends or sticky ends; sticky ends are particularly useful for recombinant work
    - sticky ends have pieces of ssDNA, can join pieces of DNA together
      - then use a ligase to join the pieces

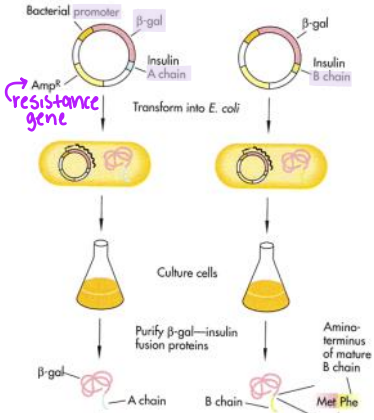


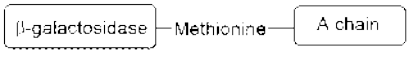




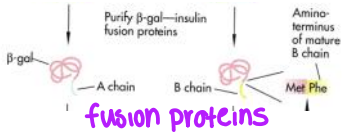


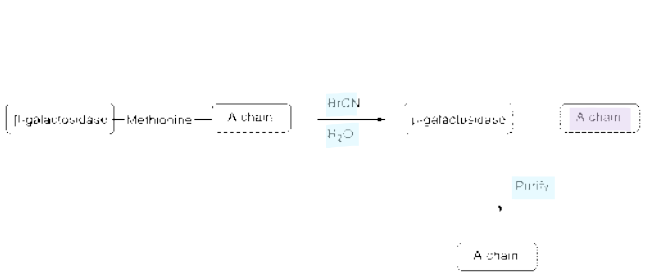
- strategy for making recombinant insulin:
  - make separate DNA for each insulin peptide (A & B chains)
  - express DNA for each in different E. Coli; created two new E. Coli subspecies
  - purify peptides separately
  - combine and form disulfide bonds to make mature, active insulin
- why make A & B peptides separately? peptides only self-assemble under special conditions
  - some purification methods with interfere with the self assembly; would cause fully assembled protein to denature itself
  - requires separately purified peptides; easier to purify each chain if prepared separately
- to easily identify transformed cells, fusion proteins were used
  - DNA for each chain was fused to gene for β-galactosidase (causes cells to become blue)
    - only blue colonies expressed the protein
  - separate plasmids were created and each plasmid was placed in different E. Coli
  - result was two fusion proteins of β-gal linked to an A/B peptide
- purification of fusion proteins:
  - homogenize
  - clarify (centrifuged)
  - ion exchange column then size exclusion column
- the two proteins of each fusion protein are separated by methionine
  - chemical reaction can be used to isolate the A & B proteins
  - treat fusion protein with BrCN in the presence of water
    - cuts chain bet methionine and A/B protein

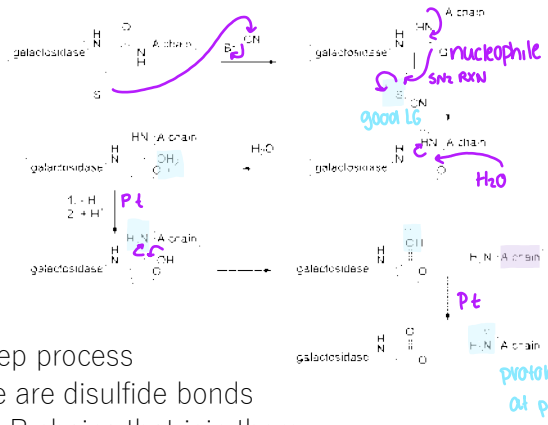




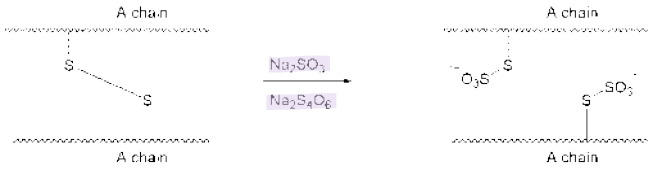
- cuts chain bet methionine and A/B protein

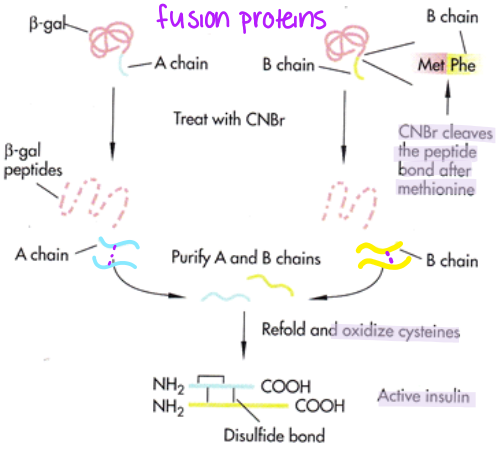




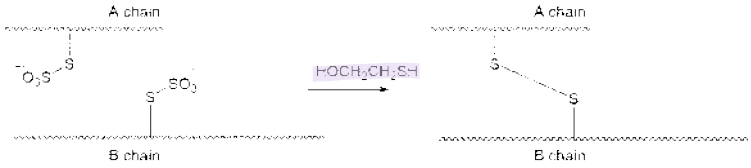


- final purification step is oxidative sulfitolysis, a two step process
  - purified A & B chains exist as pairs kinda; there are disulfide bonds formed between cysteines of two A chains/two B chains that join them
    - to break the disulfide bonds & isolate peptide chains:



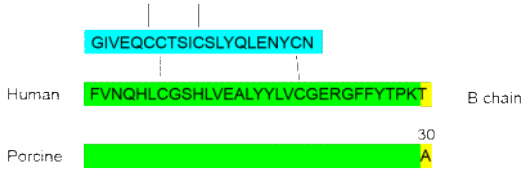


  - to get quaternary structure of insulin, mix A & B chains and form disulfide bonds:

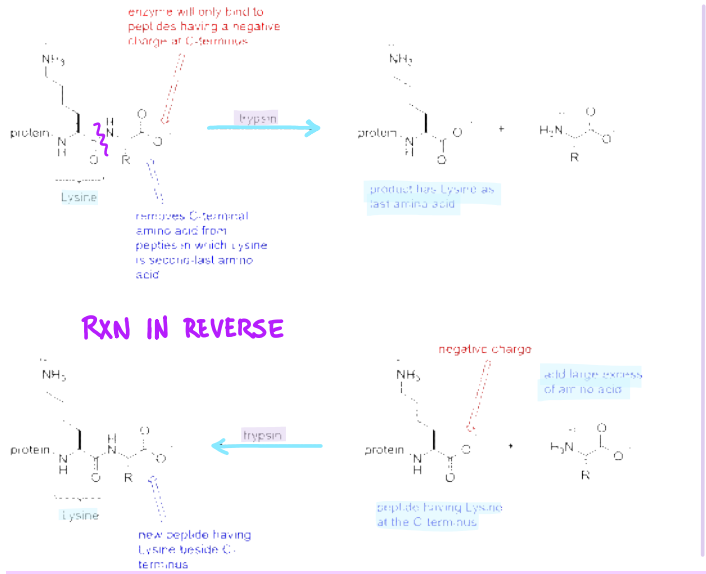


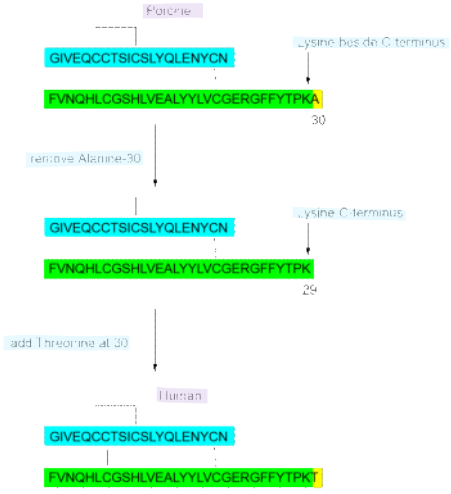
- humulin (1982) was the first recombinant drug
  - identical to native human insulin; same a.a sequence, same tertiary and quaternary structure
  - marketed as a safer form of insulin (no proinsulin)
  - how much safer is it rlly?
    - by 1982, purity of porcine insulin was very high so incidence of allergy was very low
    - porcine and human insulin both work equally well in humans; differ by only one a.a
    - semi-synthetic human insulin has also been developed; transform porcine into humans insulin
      - could be just as good as fully synthetic human insulin

SEMI-SYNTHETIC INSULIN

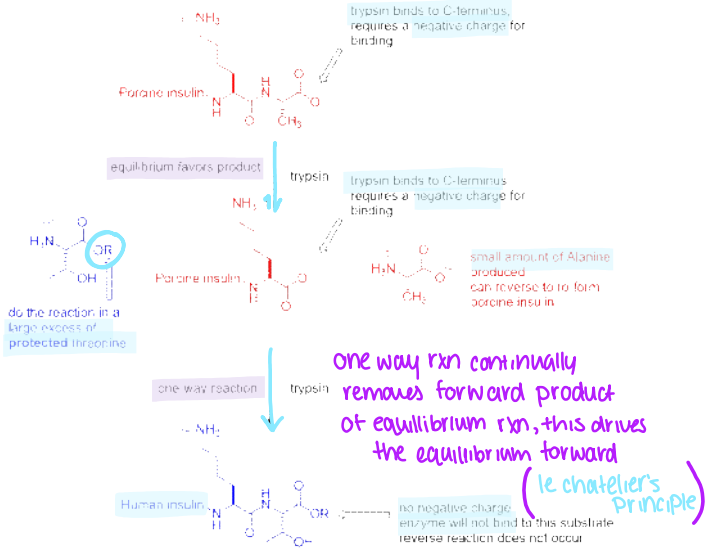


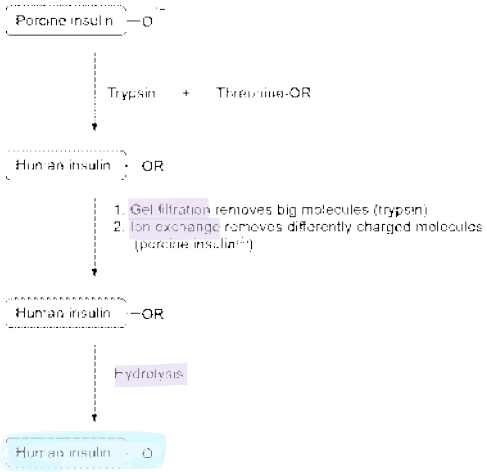
- human and porcine insulin differ by only 1 a.a in the B chain
  - position 30 is alanine in porcine and threonine in humans
- possible to convert porcine insulin into human insulin; 2 chemical rxns:
  - remove alanine-30
  - add threonine-30
- used a porcine enzyme/protease, trypsin
  - trypsin looks for proteins in which the second last a.a on the c-terminus is a lysine
    - the neg charge on the c-terminus is how the enzyme locates the c-terminus
  - trypsin then cuts off last a.a at the c-terminus
  - the product is a peptide with lysine as the last a.a. in the sequence
  - trypsin can run in reverse; looks for a protein which the last a.a on the c-terminus is a lysine
    - again locates c-terminus by its negative charge
    - if there is a large excess of an a.a acid nearby, trypsin adds it to the chain c-term
    - lysine is now the second last a.a





- trypsine can be used for both reactions to convert porcine insulin into human insulin
  - the second last a.a. on the c-terminus is lysine
  - can remove alanine and then add threonine; but need a way to differentiate between step 1 & step 2
- to control the steps, control c-terminus charge
  - made sure solution had an **excess of protected threonine**
    - threonine has a protecting group (ester) on the c-terminus; no negative charge
  - after step 1 (removing alanine), because there is an excess of threonine, trypsin adds it to the lysine
  - no negative charge left on peptide c-terminus, trypsin cant bind again
    - thus, step 2 rxn in a one way rxn
    - by **le chatelier's principle**, this forces equilibrium of step 1 rxn to favour forward product
- the protecting group also gives an easy way to purify human insulin peptide from the solution
  - use gel filtration to remove big molecules
    - trypsin is much bigger than insulin, so its removed
  - use ion exchange to remove any small amount of unreacted porcine insulin
    - human insulin is neutral bc PG and porcine insulin is negatively charged
  - hydrolysis reaction to remove protecting group



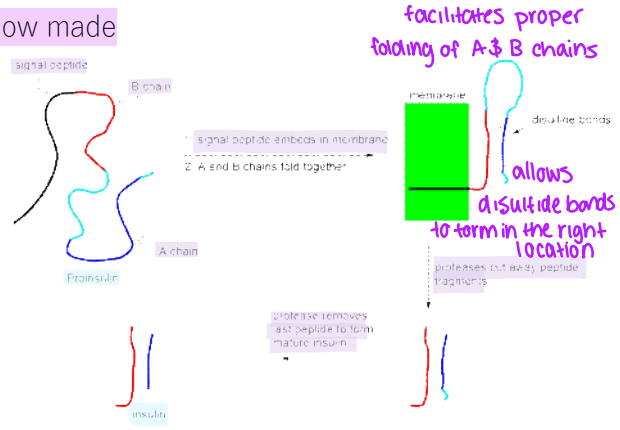


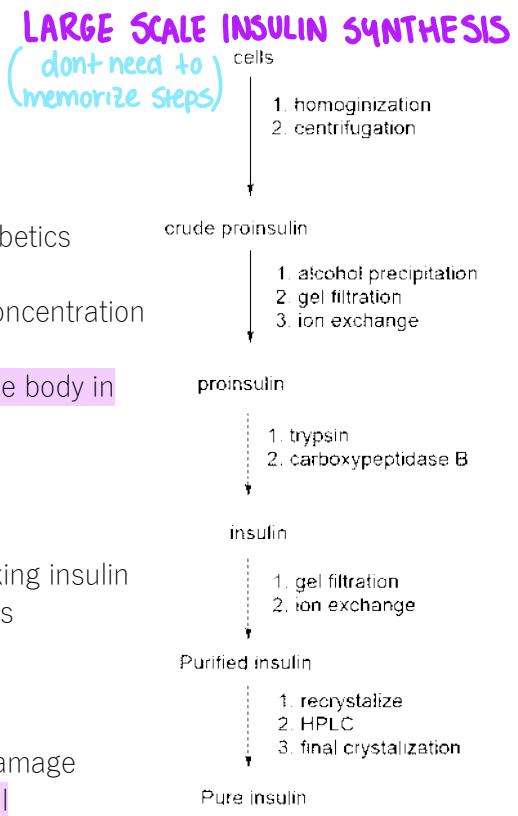
- semi-synthetic insulin is a very clean product; fixes purification and selectivity problems
- but there is a supply problem with semi-synthetic version:
  - 1 pig pancreas provides enough porcine insulin to supply 1 person for 3 days
  - 1 pig pancreas can supply semi-synthetic human insulin to 1 person for 2.5 days
    - not all porcine insulin is converted to human insulin
  - more than 170 million diabetics worldwide; so 25 billion pig pancreases required per year

- the key advantage of recombinant methods is actually supply
  - they advertised key advantage as safety bc easier for people to buy into
  - recombinant insulin is easily made in bacteria; only limited by the size of tanks its manufactured in
  - use microorganisms or plants; easy to make large quantities of insulin
  - purification is easier; fewer related proteins and can start with higher concentrations
    - can make bacteria overproduce the protein

MODERN RECOMBINANT INSULIN METHOD

- humulin recombinant insulin drug was made by making the A & B chain separately then combining the chains
  - this requires two recombinant bacteria types
- now, only need to transform one microorganism (E. Coli or Saccharomyces cerevisiae)
  - modify the method to mimic the way insulin is made in human cells
- insulin is a prohormone; a different molecule is made and then processed into the hormone
  - insulin is normally expressed in the body as a single, longer protein
  - gets cut into pieces after its folding; the single polypeptide (proinsulin) facilitates formation of quaternary structure; so using proinsulin can guarantee insulin gets folded into its proper shape
  - proinsulin contains a signal peptide; a short a.a. sequence that locates a larger protein within a cell
    - the signal protein helps to fold and assemble insulin
- proinsulin cloning and expression is how most human insulin is now made
  - animal sourced drug is rare & hard to find
  - crude proinsulin is purified & converted into insulin
- all recombinant versions of insulin today are identical to native human insulin
  - humulin (1982)
  - novolin (1991)
  - insuman (1997)
  - actrapid (2002)
- they are all identical drugs, same sequence
  - appear as different products and are regulated differently by the FDA
  - FDA is v cautious and resistant to allow generic versions of genetically engineered products
  - generic versions are now being looked at
    - called biosimilars even though they're identical

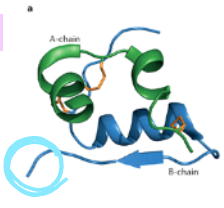




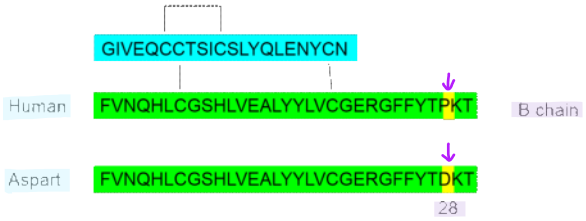
ARTIFICIAL INSULIN

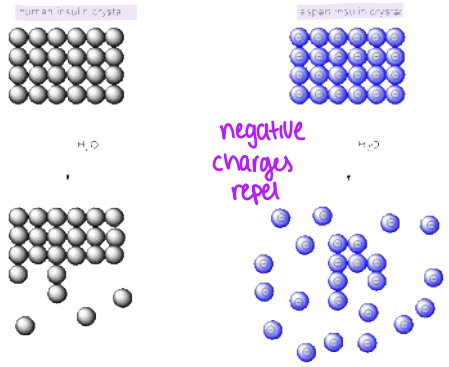
- insulin is not very drug like molecule; it's not an ideal substance for diabetics
- physiologic insulin concentrations:
  - basal levels; body maintains a constant supply of insulin at low concentration
    - approx 10^{-10} M
    - this is possible bc the insulin is produced continuously in the body in small amounts
  - transient increases when blood glucose levels rise
    - cells release insulin on demand
    - rapid response
    - levels return quickly when no longer needed; cells stop making insulin
- this created a logistical problem with the drug & taking insulin injections
  - insulin basal concentration is 10^{-10} M
  - commercial insulin drug is 10^{-3} M (more concentrated)
    - done to provide small injection volumes for the diabetic
    - & don't want to inject too much insulin into body; causes damage
  - but frequent small injections of concentrated insulin is impractical

- the design of insulin has physical/chemical properties that allow for more convenient dosing
- insulin is injected subcutaneously in fat layers below the skin
    - the drug precipitates and forms small crystal deposits in fat
    - the drug can then absorb slowly into the bloodstream
    - allows larger volume of insulin to be absorbed slowly into bloodstream
- so concentrated insulin injected in this manner allows occasional injections and a smaller insulin pen
- the crystal form helps control dissolution
    - at 10^{-3} M insulin is in oligomeric form; zinc-hexamers
    - when injected, hexamers precipitate at injection site
        - they slowly dissolve bc of crystals and provides approximate basal concentrations
- behaviour of injected insulin:
    - slow onset; have to keep track of injections
        - person must inject themselves at least 30 minutes before a meal bc blood sugar will go up
        - cannot alter mealtime once injected; need to eat what you take insulin for or else blood sugar can go too low
    - long duration
    - difficult to maintain even concentrations
        - sometimes get uneven dissolution rates
        - causes problems with high/low insulin
        - tissue damage/seizures due to this
    - injected insulin is not very drug-like
        - the natural insulin in body is continually synthesized by cells in body and released at certain concentration on demand
            - cells regulate amounts and rates of natural insulin
        - artificially injecting insulin means things don't always happen exactly the same every time
            - there is more variation in the bloodstream and basal concentration
            - its more concentrated, injection sites vary (different amounts of fat and blood supply)
            - variation in precipitation and dissolution
- insulin pumps were first developed in the 60s to counteract this non-ideal behaviour of artificial insulin
    - first developed as a backpack (mostly bc size of battery)
    - became practical in the late 90s; size of a cell phone
    - small pump with an insulin reservoir with a computer controller that controls the rate of insulin into body
    - deliver a constant supply of small amounts of insulin; mimics basal action
    - can be adjusted for faster rate/bolus delivery; short duration for meals
- why not redesign the protein? make it better than natural protein
    - its easy to change a.a sequence; can modify drug to become more drug like
    - fast acting insulin; designed for use before a meal
        - crystals dissolve quicker
    - slow (long) acting insulin; designed to mimic basal levels
        - slow, even dissolution rate
- fast acting insulin designed for quick dissolution
    - modify amino acids to speed dissolution
    - rational drug design: insulin-like growth factor 1 (IFC-1) has similar a.a sequence to insulin
        - order of Pro(B28) and Lys(B29) reversed compared to insulin sequence
        - proteins tends not to oligomerize & dissolves quickly
    - switch the order of a.a at positions B28 and B29 in insulin to make it dissolve faster
        - small a.a. change at end of chain, doesn't otherwise affect structure or function
        - called lispro, designed in 1996
    - lispro does not self associate well, zinc crystals break up quickly
        - dissolves 300X faster than regular injected human insulin
        - good for bolus injections before meals

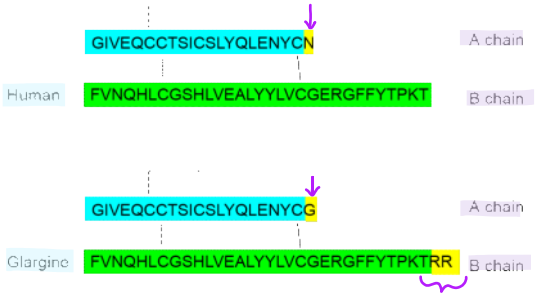


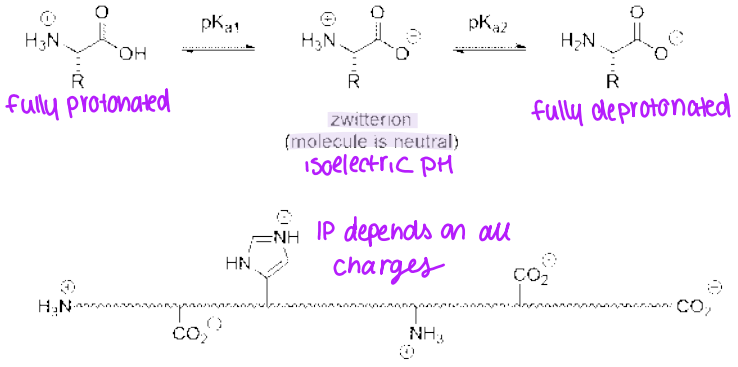
- aspart (2017); proline at position B28 replaced by aspartic acid
  - creates charge repulsion between molecules in the hexamer crystal (crystal is less stable)
    - proline is neutral while aspartic side chains negative at pH 7.4
  - crystal dissolves quickly, good for bolus injections before meals



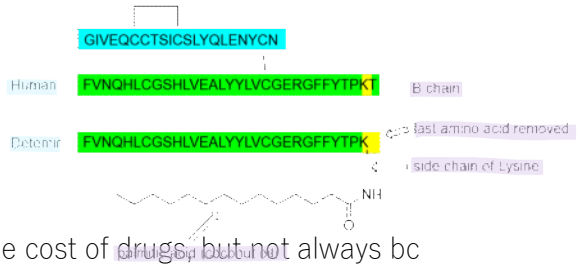


- diabetic experiences large variation in basal insulin levels
  - creates metabolic problems, protein damage, discomfort (if insulin low), frequent insulin injections to mediate basal levels
  - want even basal levels like non-diabetics
- long-lasting insulin; slow dissolution increases time between injections
  - less variation in blood amounts
  - achieved by lower solubility; crystals take longer to dissolve
  - glargine (2000); two arginines added to B chain (now 32 amino acids)
    - changes isoelectric pH of protein from 5.4 to 7.2
      - lowers the solubility at pH 7.4 bc more likely to be neutral at that pH
    - isoelectric point: pH at which overall charge on a zwitterionic molecule is zero (neutral and insoluble)
      - in proteins, IP is determined by pKa's & charges of N-terminus, C-terminus & side chain groups
    - but adding arginines made the molecule unstable; to fix this, modified a residue on the A chain
      - asparagine at N19 on A chain was replaced by glycine to stabilize quaternary structure of protein





  - detemir (2006) employs fatty acid attachment
    - added a lipophilic side chain (fatty acid)
      - attached palmitic acid (comes from coconut oil) onto lysine of B chain
      - also removed threonine (last a.a. of B chain) to reduce problems associated with attachment of fatty acid
    - lowers solubility in water; slower dissolution in blood
    - drug is "soluble" in subcutaneous fat
      - slow release and even rate of dissolution into blood



- advantages of genetic manufacturing:
  - less risk of contamination (allergy, viruses)
  - large quantities easily available
    - this lowers the cost of manufacturing and should lower the cost of drugs but not always bc companies want more money. if a genetically manufactured drug isn't cheaper, its bc of business reasons not scientific reasons
  - possible to modify properties of proteins to make them more drug-like and user friendly

TOPIC 9: HUMAN GROWTH HORMONE

HGH INTRO

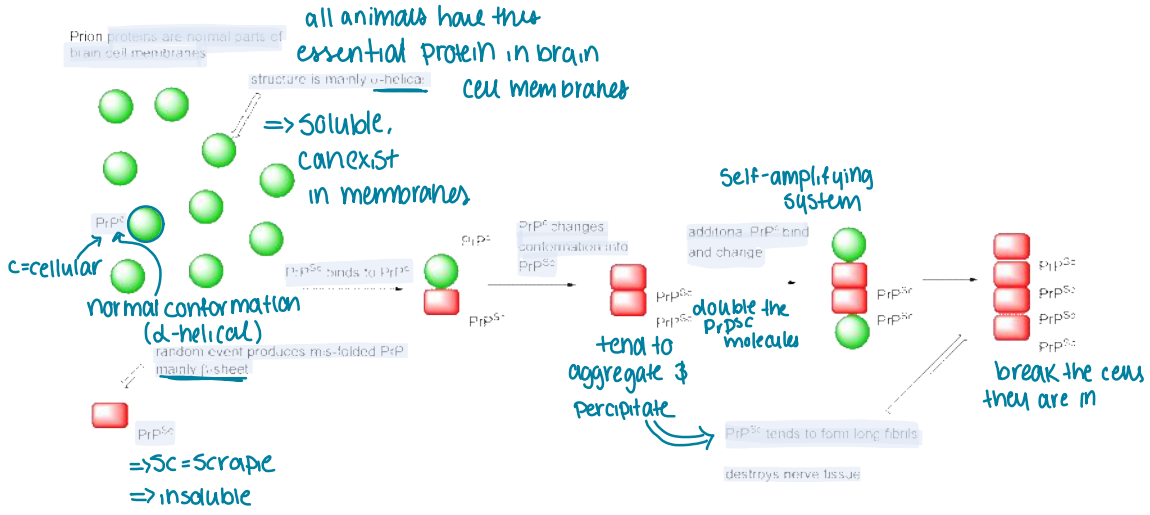
- human growth hormone aka somatotropin is a 191 amino acid protein (~4x the size of insulin)
  - secreted by the pituitary gland
  - stimulates growth & cell regeneration
  - impo for development & damage repair
  - secreted continually throughout life (400ug/day in adults)
  - larger amounts produced during growth (700ug/day in adolescents)
- HGH deficiency aka hypopituitary dwarfism: a rare condition in which the pituitary doesn't produce enough HGH
  - short stature, short digits, limited elbow mobility, lower back issues
- treatment is with hGH
  - animal source GH doesn't work on humans, need human-sourced HGH
  - isolate and purify HGH from human cadavers
    - bc the hormone is prod in very small amounts in our body, need to get the whole pituitary gland as a source; only possible if you take from dead bodies
  - cadaver-derived hGH available in the late 50s
    - people donated their body to science
  - very limited supply, only approved use was hypopituitary dwarfism
    - only source was HGH in body at time of death
  - approx 29000 children treated worldwide bet 1958-1985
- to harvest HGH from cadavers:
  - remove brain and cut out pituitary; common to include small amount of brain tissue
    - want to get every bit of the gland possible
  - requires ~8 cadavers/year/child
    - weekly injections, must be continued for 10-12 years
- impossible to purify anything 100%; proteins from living sources always contain small amounts of contamination
  - proteins have v similar properties, hard to purify
  - proteins
    - may cause allergy
    - enzymes (shelf life)
    - other functioning proteins (various effects)
  - viruses; may cause infectious disease
    - purification include special anti-viral steps; today this risk is negligible but not zero

CADAVER-DERIVED HGH & C-J DISEASE

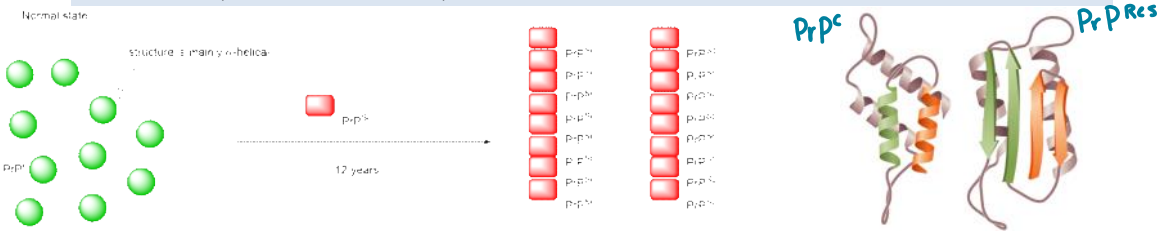
- in 1985, there were 4 cases of an extremely rare brain disease called creutzfeldt-jakob in hGH patients
  - disease turns brain into spongy material and slowly destroys brain tissue
  - the disease takes more than 12 years to develop; very slow acting disease
  - the four people had all been given cadaver-hGH during the 70s
  - the use of cadaver-derived hGH was suspended in 1985 bc of this
- creutzfeldt-jakob is in a family of diseases called transmissible spongiform encephalopathies
  - 100% fatal brain disorders; death ~1 year after diagnosis, no treatment
  - dementia, blindness, involuntary movements, strange behaviours
    - ex of strange behaviour in similar disease in sheep is scrapie
      - sheep will scrape their bodies against rocks, trees etc and scrape their skin off
  - characterized by formation of amyloid plaques (clumps of protein that accumulate) that destroy the brain tissue; brain will deflate/collapse on itself
  - most cases arise spontaneously (more than 80%) and before the 70s thought that genetics account for the remainder

- rare diseases in various organisms:
  - bovine spongiform encephalopathy; mad cow disease
  - scrapie (sheep)
  - kuru (human cannibals)
  - creutzfeldt-jakob (humans)
    - spontaneous rate is 1 case in >10M people (v rare); average age of onset is 60+ years
- mad cow disease, scrapie and kuru have patterns that suggest infectious component

- does C-J have an infectious component?
  - multiple cases of scrapie in a flock of sheep
    - the disease is v rare, the fact that multiple sheep in a flock get it suggests an infectious component
  - multiple cases of kuru in a village happen in waves (when one person gets it, multiple other people do)
  - the diseases appeared to have very long incubation times, so they were thought to be caused by "slow viruses"
  - research in the 70s indicated the infectious agent was a single protein
    - not virus, bacteria etc
- the infectious agent is a single protein found in the brain of the organism
  - when a protein becomes infectious, its called a prion
- hGH isolated from cadavers and C-J
  - to date, more than 230 cases of C-J in patients given cadaver-sourced hGH
    - 29k given cadaver-derived hGH; 1 case in every 126 patients
    - onset in middle age
  - normal rate of C-J in humans is 1 case in >10 M and later onset
- prions are an infectious conformation of a protein; misfolded proteins with the ability to transmit their misfolded shape onto normal variants of the same protein



  - if one of the proteins mistakenly takes on a β-sheet conformation, eventually all other proteins will be converted and the fibrils will destroy brain tissue
  - the amino acid sequences of the normal & infectious variants are identical, just the 3D conformations differ
  - PrP^{C} aka prion protein cellular is the normal variant of the protein found in cells
  - PrP^{Sc} aka prion protein scrapie is the infectious variant in sheep with scrapie
  - PrP^{Res} aka prion protein research is a form made for research purposes
    - very similar to PrP^{Sc} but not exactly the same
    - when the protein is in the β-sheet form, its v hard to work with bc insoluble & tends to aggregate
    - PrP^{Res} is a special form of the protein created to make it easier to work with in the lab

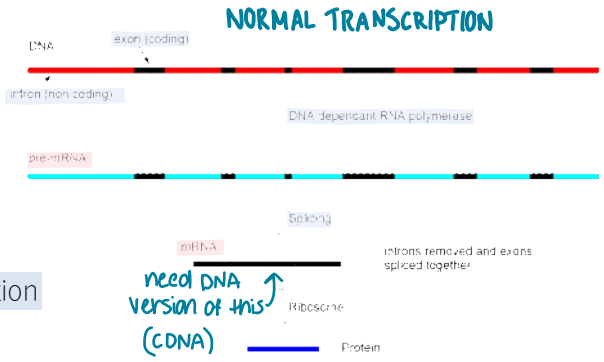


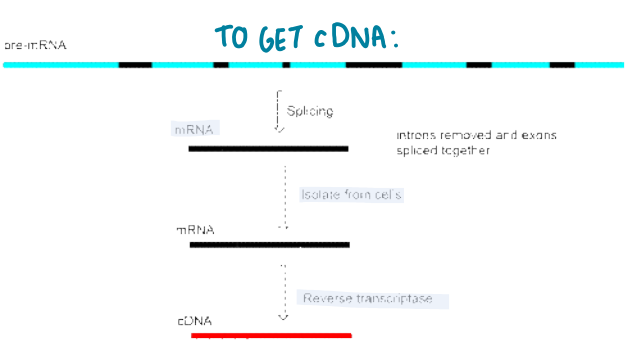
    - prions arise from conformational changes
        - one molecule in α conformation randomly changes to β conformation
            - β form binds to α form and cause them to change conformation to β
            - exponential growth in β forms
            - over time, β fibrils build up and disrupt nerve cell structure
- PrP^{c} is a normal component of nerve tissue found in neurons & glial cells
    - membrane bound protein
    - the sequence is highly conserved (similar) across species
        - this is an indicator that its an impo protein for normal functioning of our brains
    - the function is not fully understood, but seems to have a role in cell adhesion, apoptosis and the cellular stress system
- the disease arising spontaneously is a rare event
    - statistically unlikely in younger animals bc takes time for random chance to give rise to a prion protein so its mostly in older animals/humans
    - develops slowly (12+ years in humans)
- PrP^{Sc} is a transmissible molecule; the disease is infectious by eating nerve tissue (brain) from the **same species**
    - β sheet form is very robust, it doesn't get fully digested in the stomach
        - is able to pass into the body (not fully understood how)
    - very long incubation period; several years in sheep or cows and decades in humans (kuru)
- **originally thought to be non-transmissible across species** but this changed with an event during 80s and 90s
- mad cow disease is very rare (so rare the exact rates aren't known; likely less than 1 in 1M)
    - affects cows 4-5 years old
    - arises spontaneously; affects single animals
    - herd outbreaks are v rare, wondered if they were even possible
- the meat industry thought of using meat and bone meal as a money-saving way to add protein supplement in animal feed
    - used **offal in cattle feed**; offal is the part of cow that humans don't eat (waste)
        - internal organs, entrails, **brains**
        - small market for these products as human food
        - sold it to make cattle feed
        - increase cow's food supply to make them grow faster
    - so **cows were eating cow brains**; can have exposure to PrP^{Sc} & be infected
- also "downer" cattle were cows that were unable to stand up and normally suffer so they were normally killed
    - until 2003 there were few regulations on the use of the carcass of these dead cows that were sick
    - they were sold for use in different products (**including brain tissue**)
        - rendered (basically ground up the entire body) for use in pet food and cattle feed
        - butchered for human food
- there was an outbreak of mad cow disease in europe from 1986-2001 (used to be v rare disease)
    - england had 400 cases in 1986 and 120k cases in 1993
        - more than 180k cattle were affected
    - france; more than 300k affected
    - over 4.5 M animals were destroyed (carcases incinerated) to counteract this and get rid of the infectious agent (prion)
- mad cow disease and C-J:
    - initially thought to be unable to jump species bc human and cow proteins are different
    - but the first case of **variant C-J** was observed in 1995
        - prion protein in β sheet was slightly different and symptoms diff (brain looked slightly different)
        - this was traced back to the person **eating beef from cows with undiagnosed mad cow disease**
    - to date, more than 200 people in europe affected by eating meat from infected cows

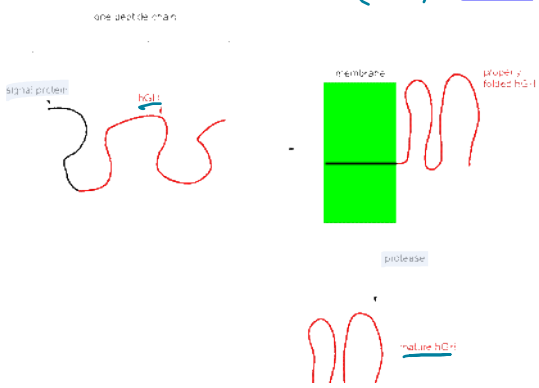
- north american response
  - 1997; US banned use of meat and bone meal in cattle feed
    - still allowed butchering of downer cattle
  - 2003; first case of mad cow disease in canada
  - 2004; US banned use of cow blood in cattle feed
  - took until 2003 to update regulations on downer cattle
    - banned slaughter for human use
- rise of zombie deer in NA; chronic wasting disease
  - an infectious spongiform encephalopathy (prion) in deer
  - first seen in captive deer in the 60s and in wild deer in 1981
    - captive deer were also fed animal feed from downer cattle/deer
  - found in most of western US by the 90s and canada in the 2000s
  - recently found in wild moose, reindeer, elk
  - rates in captive deer now as high as 79%
- C-J and hGH:
  - pituitary located in center of the brain & harvest likely included brain tissue
  - PrP^{C} is a normal component of brain tissue so there can be small contamination with PrPSc from cadaver
    - most medical cadavers from older people & older people are more likely to have PrP^{Sc}
  - pituitaries are processed in batches so one diseased pituitary can contaminate an entire batch of hGH
    - each drug batch is shared by many patients

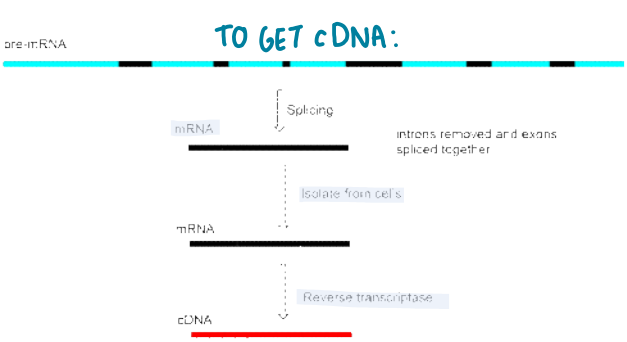
RECOMBINANT HGH

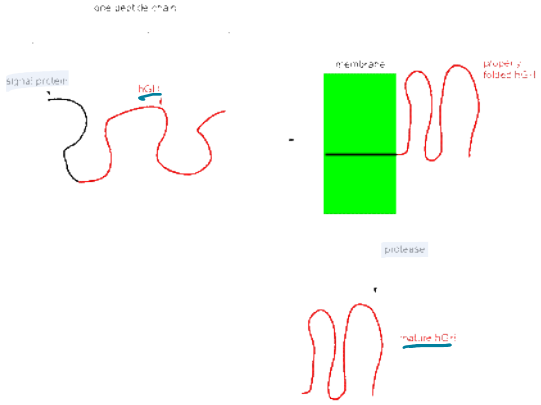
- animal sources don't work for hGH
  - hormone is highly species specific, need human version
- cadaver sources no longer possible bc risk of prion contamination
- by genetic engineering, can make human version with no risk of contamination & can supply large amounts
- human growth hormone is a single protein chain 191 a.a long
  - DNA was too long (~600 nucleotides) for machine synthesis at the time (1981)
- need cDNA to transform bacteria
  - cDNA = coding DNA (exons only, introns removed)
- isolate mRNA from cells and use reverse transcriptase to make cDNA
- isolated mRNA from pituitary tumors (get via endoscope up the nose)
  - normal pituitary cells will not grow outside the body but tumor cells sometimes do
  - tumors can provide enough cells for isolation
- HGH is normally expressed as a pro-protein; fusion protein of HGH and a small signal protein required for proper folding
  - signal protein causes pre-HGH to be transported to the membrane and be embedded in the membrane
  - this allows HGH to fold into its proper shape
  - proteases cut the protein off when its folded
- so in order to fold the protein properly, need the signal protein function



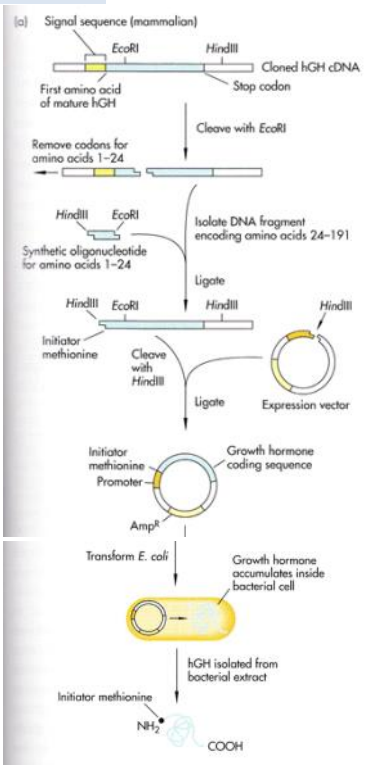




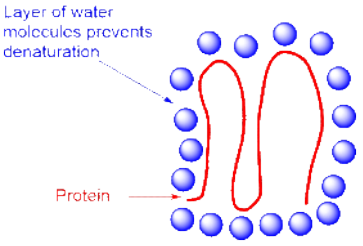

pre-mRNA
TO GET cDNA:
Splicing
mRNA
introns removed and exons spliced together
Isolate from cells
mRNA
Reverse transcriptase
cDNA


membrane
protease

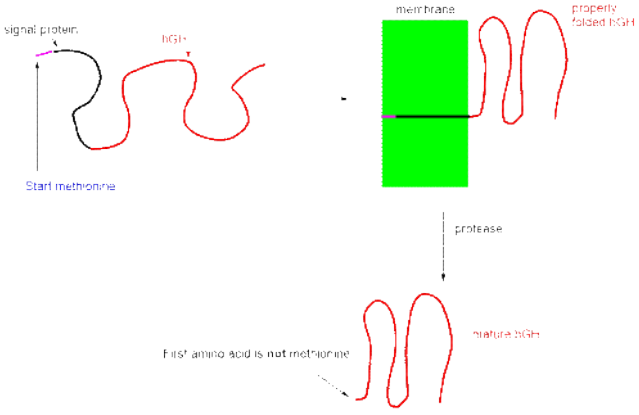
- human signal protein doesn't work in bacterial cells
  - needed to remove human signal protein sequence from cDNA for bacterial expression
  - used a restriction site in middle of gene
  - but in the 80s when this was done, there was a limited number of restriction enzymes available
  - closest they could get to the site between signal protein and HGH was inside the coding for HGH itself
  - lost codons for first 24 a.a's of HGH
  - replaced with synthetic DNA of first 24 a.a's of HGH only
  - added extra Met at N-terminus of synthetic DNA
    - needed starting point for the bacterial ribosome to read the gene



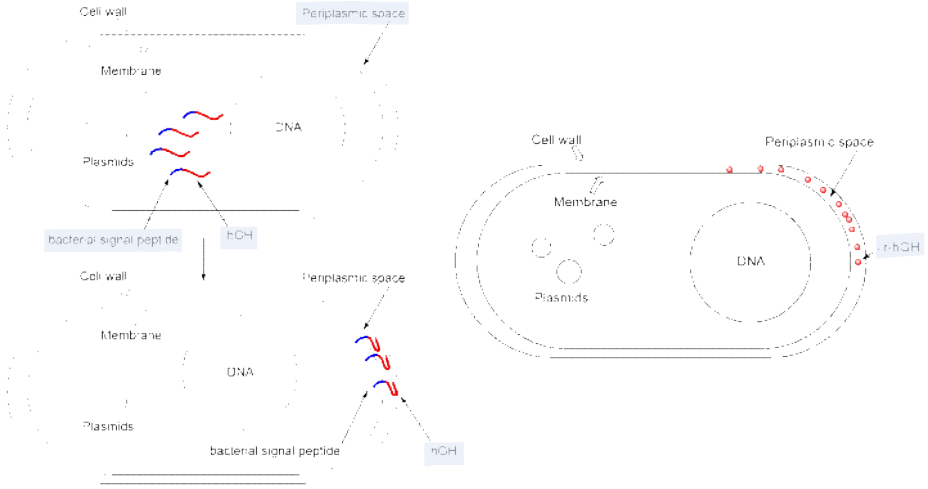
- the DNA was then put into a plasmid that was used to transform bacteria and make the synthetic version of HGH
- the protein contained an extra methionine
  - hGH has 191 a.a
  - r-met-hGH had 192 a.a
- general purification: (don't need to memorize this sequence)
  - lyse bacteria (sonication)
  - clarify
  - precipitate
  - ion exclusion then size exclusion
  - final precipitation
  - re-dissolve and lyophilize (freeze drying)
- lyophilization; method of preserving protein for long term storage (longer shelf life)
  - get less chemical rxns in solid phases bc solids are usually more stable than liquids
  - freeze sample, and while frozen remove H_2O using reduced pressure (vacuum)
    - water sublimes and leaves behind a fine powder
    - preserves the 3D structure of the protein
  - maintains tertiary structure bc layer of water molecules on surface of protein not removed
    - prevents denaturation of the protein, resists unfolding
  - can store drug as a powder

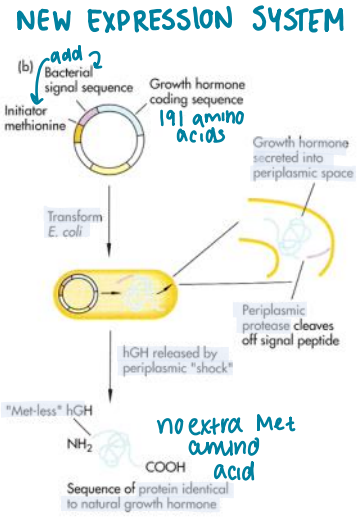


- why the extra methionine? bacteria needs a start codon to start expressing the gene
  - mRNA contains a start codon for ribosome attachment
  - start codon is AUG which codes for methionine
    - first a.a in all expressed proteins (eukaryotes & prokaryotes) is methionine
  - hGH is normally expressed as a fusion protein, so Met is not in the final peptide
    - starting methionine is located on the signal peptide & is lost during protein maturation
- normal production of hGH in humans:
  - initial peptide has Met as first a.a
  - final product does not contain Met at position 1
- r-met-hGH (protropin) marketed in 1986
  - almost identical to hGH
    - one extra a.a, still effective
  - safe; no prion risk (this was a breakthrough)
  - no longer sold tho, better products now
- the better products address difficult batch purification
  - protein was produced inside bacteria
  - lots of other proteins present after cell lysis
  - difficult to isolate the desired protein, more $$$ product



- for easier purification, design a protein to be secreted
  - protein is produced inside the membrane and secreted outside the bacteria
    - less bacterial contamination
- protein secretion:
  - bacterial signal peptide (fusion protein) used to carry r-hGH outside the membrane
  - made genetically modified bacteria to produce HGH fused to a bacterial signal peptide
    - the methionine is located at the beginning of the signal peptide (start of replication)
  - the bacterial peptides causes the protein to be transported to the outside of the cell
  - the signal peptide becomes embedded in the membrane such that the HGH is in the periplasmic space: the space between the cell membrane and the cell wall
  - after signal peptide is cut off by bacterial enzymes, r-hGH is left in the periplasmic space
  - to get access to HGH, fracture the cell wall but keep the membrane intact
    - get HGH in isolated state without all other bacterial proteins

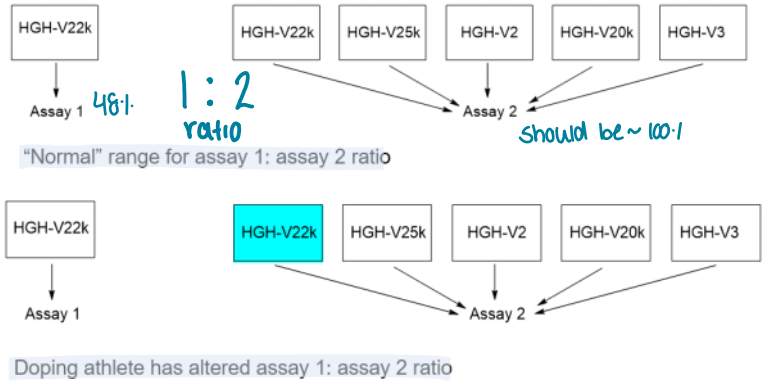




  - cell wall was fractured using osmotic shock:
    - incubate cells in 0.5M sucrose
      - hypertonic, sucks water out of cells
    - remove medium and replace with cold dilute salt solution
      - hypotonic, cause water to rush inside cell
      - this fragments the cell wall
  - purification steps include lots of modern chromatography & lyophilization
- protein secretion is used often nowadays in manufacture
- two general manufacturing methods for recombinants:
  - batch: grow organism in large batch then pool and purify protein
  - continuous flow: organism is suspended in a medium that continuously circulates
    - organism continuously secretes protein into the medium. collect secretions and purify
- recombinant hGH: nutropin, humatrope, genotropin, norditropin, saizen
  - all identical to HGH; same a.a sequence and tertiary structure
  - different name brand versions of the same thing instead of generic version bc FDA is v cautious about licensing these
- now that there is a large supply of hGH, more uses
  - original use was for hypopituitary dwarfism in children
  - hGH has therapeutic potential in other conditions
    - adult growth hormone deficiency
    - wasting conditions (lose lots of weight): AIDS, cancer, immobility
    - improve resistance to injury in seniors
    - faster wound healing

- lots of internet scams with hGH in pill form
  - hGH must be injected, it's a large protein
- illegal use of r-hGH; performance enhancement in sports
  - normally: 400ug/day in adults and 700ug/day in adolescents
  - some athletes take >1000ug/day
  - first reported use in 1981 (cadaver-sourced hGH)
  - banned by olympics in 1989
    - but testing for it is impossible bc the sequence of drug is identical to normal HGH
  - first effective testing in 2004, and first athletes testing positive in 2010
  - so far 15 athletes caught
- effects of hGH on athletes is unknown, lack of clinical data. this is what we know from seniors:
  - increases lean body mass; ratio of muscle to fat
    - effect on athletic performance? unknown
  - increased injury resistance but effect very small
  - no evidence for increased muscle mass
  - any benefits are likely small bc hGH by itself gives no benefit
- detection of hGH in sports:
  - impossible to detect directly bc r-hGH identical to hGH
  - indirect methods must be used: measure total amounts and ratios of isoforms of hGH
    - human pituitary produces 5 isoforms of hGH; r-hGH is identical to HGH-V22k (the main isoform; normal abundance is 48%)
    - measure amounts of all 5 forms and look for abundance of the main form (should be 48%)
    - total amount are also important
    - long term use gives back-regulation
      - disturbs the ratio in other ways
      - get different ratios of other isoforms

Isoform	Length	Characteristics
HGH-V22k	191aa	Main isoform normally 48!, drug is this isoform
HGH-V25k	191aa	Glycosylated version of V22k
HGH-V2	230aa	Retains one intron
HGH-V20k	176aa	Deletion of amino acids 32-46
HGH-V3	219aa	Alternate processing of one intron

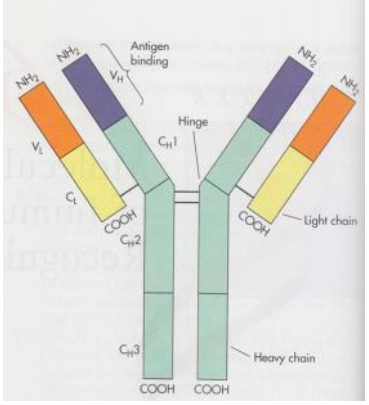


TOPIC 10: VACCINES; HEPATITIS

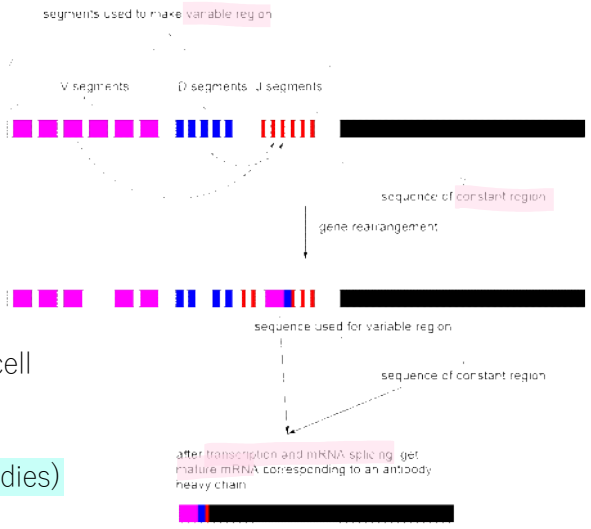
December 3, 2020 11:56 PM

VACCINES

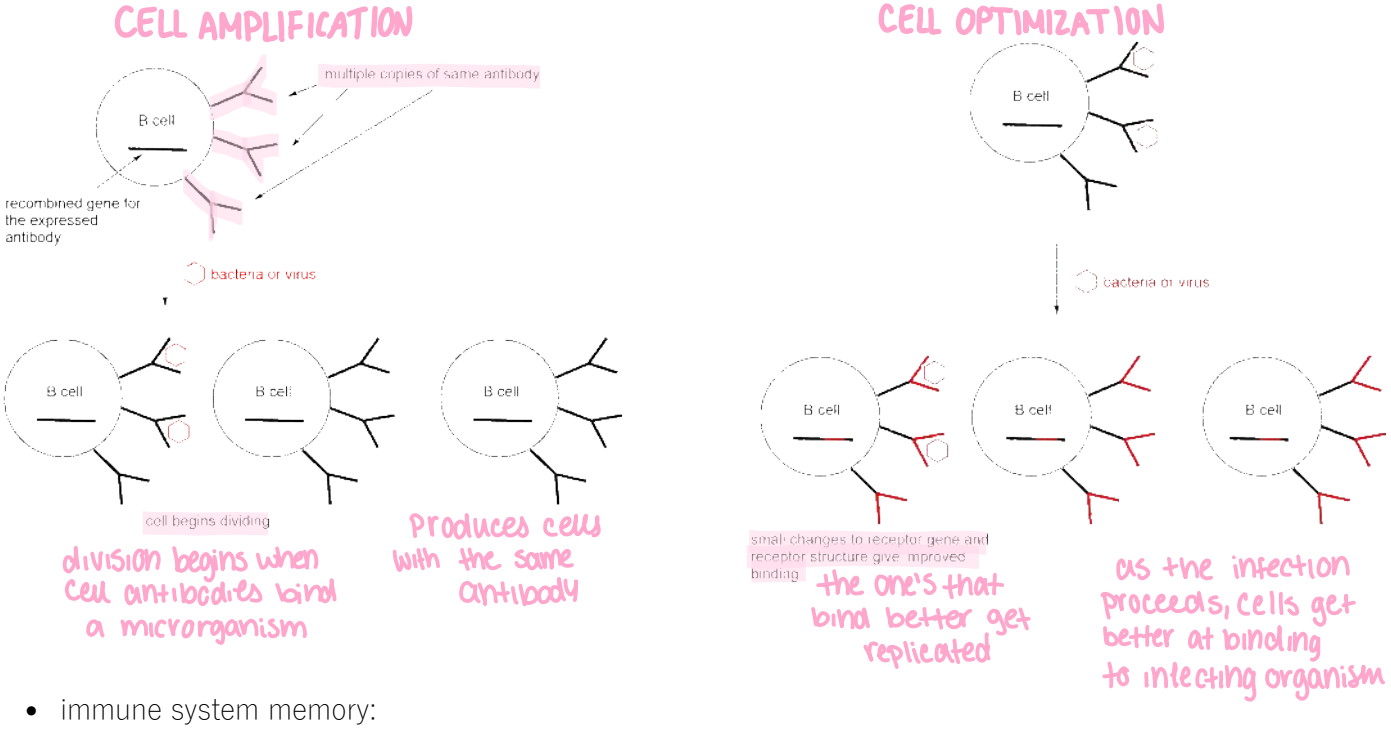
- vaccines prevent illness; better to avoid a problem than to try and fix it
- our immune system retains a memory of infection
- many organisms can only infect once
  - infecting microbes are killed before numbers become large
- vaccines are drugs that produce artificial immunity (memory)
- subsequent exposure produces fast immune response
  - memory of an infection without the infection
  - gives immunity or a less severe, short infection
- immune system "design considerations":
  - system must have a way to identify when infection occurs
    - to know when something is worth defending against
  - impossible to know beforehand what will infect
    - system must have a way to adapt, must be flexible and adapt how it operates
  - system must respond before the host animal dies
    - deadly organisms usually kill host before immune system can respond to it
  - if a microbe infects once, it is likely to infect again
    - its useful to have a "memory" system
- in the immune system, nobody is in charge (there is no central coordinating system)
  - body/immune system is a collection of cells
    - communication is accomplished using chemicals
  - functions using distributed control system
  - ant hill analogy, no one tells them what to do, chemical communication what needs to be done
- adaptive immune system general operation:
  - B cells and T cells
    - each has unique receptor or antibody on the outside of cell
    - receptor/antibody binds to small protein fragments (several a.a. long) to be able to carry out its function
  - receptor & antibody structures are made semi-randomly
    - done to conserve space in genome
  - each cell "shuffles" special gene segments randomly
  - produces large structural variation in receptors
  - each cell produces a single receptor structure
    - each T/B cell has its own gene made by randomly shuffling gene segments
  - with this random assembling, its possible to make antibodies and receptors that bind to our normal body proteins; this is bad, attacks our protein
  - so our body removes any receptors which recognize host cells or molecules
- antibody structure:
  - Y shaped molecule
  - has conserved/constant (C) regions and variable (V) regions
  - antibodies are made by combining constant and variable genes
  - each antibody has two different chains of protein
    - the light chain and the heavy chain
    - N-terminus at the tips of the Y (top)
    - C-terminus at the stem of the Y (bottom)
  - the variable part at tip is the part that will bind to molecules it recognizes



- gene shuffling creates different antibodies and receptors
  - process is random
  - each cell expresses a different antibody mRNA
  - each cell carries a different antibody
  - the gene for antibodies has variable segments & a constant region
  - can make a variety of genes by combining one gene from each variable segment with the constant region and expressing this unique mRNA

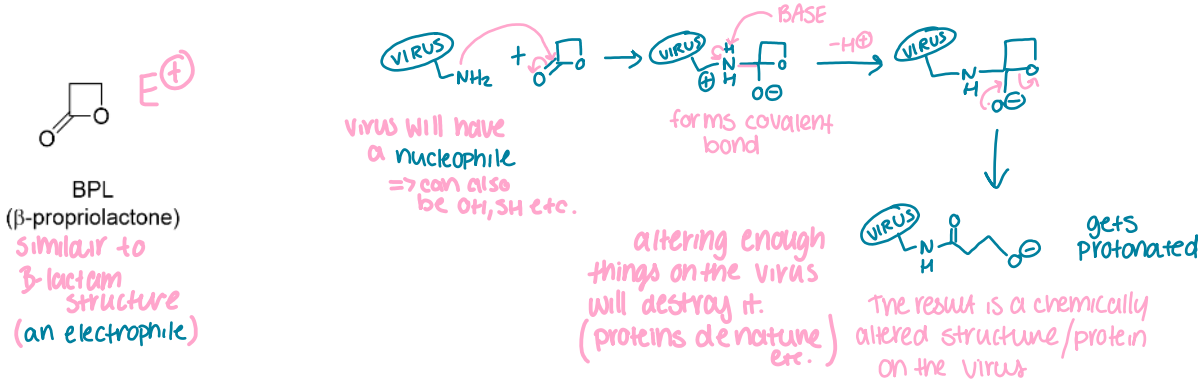


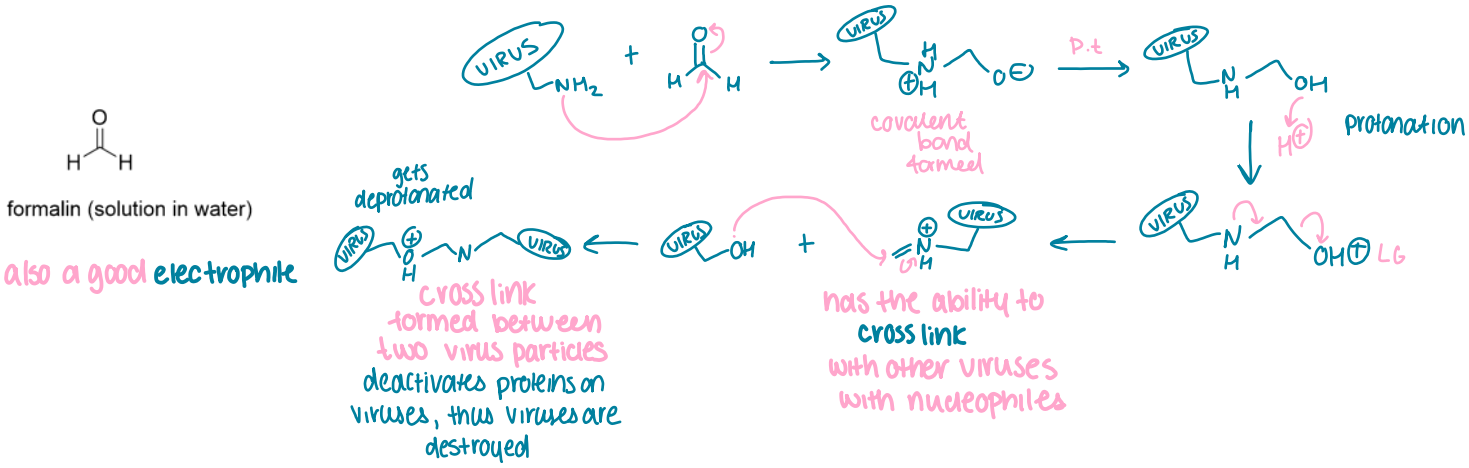
- this gene shuffling occurs at an early developmental stage of the cell when its being made. this is now permanent for the particular cell
- response during infection:
  - only a small number of B or T cells have receptors (or antibodies) which bind to the invader (50 to 100 cells)
  - when these cells recognize that they're binding to a microorganism, they divide
  - as they divide, they make small changes to the gene producing their receptor
    - receptor binding is enhanced
  - result is more cells that can fight invader, and each generation of these cells works better
  - this process takes several days, illness lasts a few days before you start getting better



- immune system memory:
  - after infection, most of the responding B/T cells die
  - but a population of the cells is retained as memory cells
    - higher numbers than at the start of next infection
    - have "optimized" receptor binding
  - provides an immediate supply of the appropriate recognition molecule during subsequent infection
- principle of vaccination: need to stimulate immune system to produce memory cells
  - higher copy numbers with optimized binding
  - create a "stimulated" infection; trick immune system to create a memory without illness

- stimulated infection methods:
  - attenuated virus or bacteria
    - chemical treatment to remove virulence elements (part that makes it deadly)
      - still has same structure as the dangerous organism
    - produces mild illness
    - once immune system clears the infection, memory cells are retained
  - dead virus or bacteria
    - chemical treatment to kill microbe; as long as the chemicals that immune system responds to are intact, memory cells will be made
    - does not produce illness
      - fever, soreness
    - stimulate the production of memory cells
      - body responds to foreign proteins on dead organism
    - because this type of vaccine doesn't create an infection, there is less organism in the body that with an active illness
      - requires booster shots to "pump up" the number of memory cells
- vaccine manufacture ex: flu vaccine
  - eggs are inoculated with a viral culture
  - incubation for 72 hours; virus grows inside the eggs
  - cooling; stops viral growth
  - harvest eggs
    - shells opened
    - allantoic fluid removed (fluid in small sac near yolk)
      - only harvest this liquid
  - clarification; want to isolate solid
    - centrifuge
    - collect fraction 75-121 kDa (contains virus)
  - now have virus in live form, but don't want live form for the vaccine
  - purification:
    - inactivation with chemical treatment (kill virus)
      - BPL (β-propriolactone)
      - formalin (formaldehyde in water)
    - dialysis; removes extra proteins and reagents added
    - splitting; detergent is added to disrupt viral structure
      - triton or ether
      - now have a "double-dead" virus
    - ion exchange
    - formulation; involves adding Ca^{+2} and Mg^{+2} to stabilize tertiary structure of flu proteins
- major side effect of flu vaccination is due to egg allergy
  - bc vaccine is grown in eggs, & impossible to remove all egg protein
- vaccine inactivation mechanism:





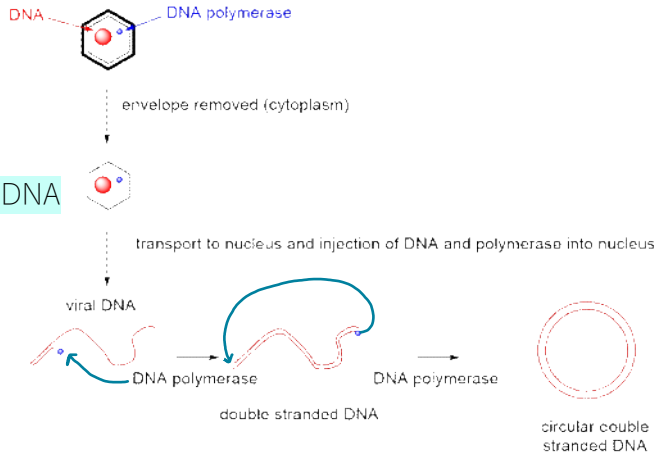
HEPATITIS

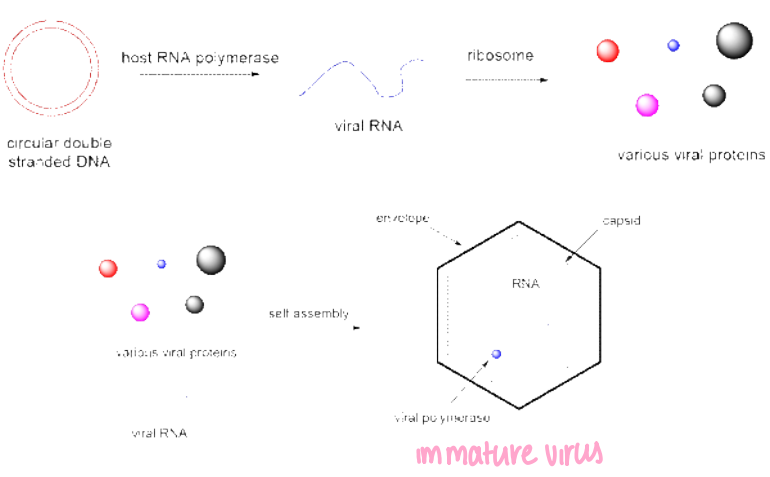
- disease of liver (inflammation, jaundice); can lead to cirrhosis, cancer, death
- infectious hepatitis (spread by contact)
  - type A; virus discovered in the 40s
- serum hepatitis (direct exchange of body fluids)
  - type B; virus discovered in 70s
  - type C (non-A-non-B); virus discovered in 90s

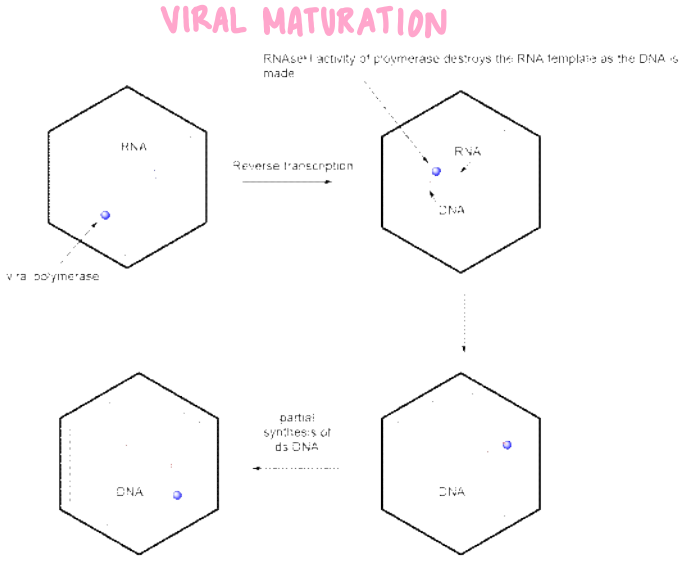
HEPATITIS B VACCINE

- viral infection of liver
- generally **acute infections** in adults
  - usually no long term damage
- **chronic infections** (children and some adults)
  - asymptomatic; slow liver damage
  - eventually develop cirrhosis, cancer, death
- transmitted by body fluid
  - congenital infections
    - common for children; child gets infected by mother during childbirth
  - sexual transmission
  - illegal intravenous drug use
  - tattoos, piercings (rare)
- common infection in the world; about 1/3 of pop will be infected at some point in life
- more than 300M chronic infections, and ~130M new infections per year
- most cases in africa, asia, SA; uncommon in developed world <1%
- discovery of virus
  - antigen protein discovered in blood (1965)
    - antigen = something that binds to an antibody
    - associated it with hep B infection
    - later shown to be aggregates of viral envelope protein
      - the envelope protein usually is embedded in viral membrane
      - protein on outside of an unmatured virus can aggregate
    - hepatitis B surface antigen (HbsAg)
  - virus discovered in 1970 & genome sequenced in 1980

- structure of hep B virus:
  - 8 genotypes; variable distribution around world
  - DNA; partially double stranded (mostly ss with short piece of ds)
  - DNA polymerase
    - a reverse transcriptase
    - has an RNAse H (destroys template RNA)
    - not as sloppy as HIV pol
  - protein X; unknown function when discovered
    - now know it has multiple functions for replication of the virus
- life cycle of hep B virus:
  - taken up into **hepatocytes** (liver cells)
  - viral envelope removed in cytoplasm; exposes capsid
  - capsid transported into nucleus
  - DNA & DNA pol released into nucleus
  - viral DNA pol completes the conversion of ssDNA into dsDNA
  - viral pol circularizes the DNA; plasmid-like structure
- replication of hep B virus:
  - host RNA polymerase transcribes various viral RNAs
  - viral mRNA translated into viral proteins
    - HBsAg is highly overexpressed
      - ends up embedded in surface of hepatocyte
      - some also leaks into bloodstream
  - viral proteins self-assemble around viral RNA and viral pol
  - viral DNA pol reverse-transcribes RNA into a single strand of DNA
  - viral DNA pol makes part of the second DNA strand (makes small dsDNA part)
  - virus particles secreted by the cell, steal some host membrane to become the viral envelope



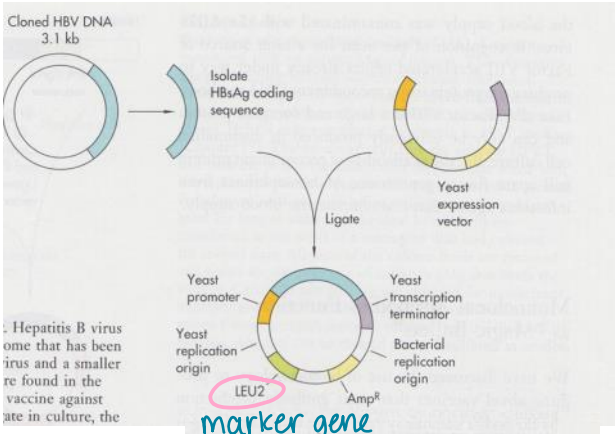




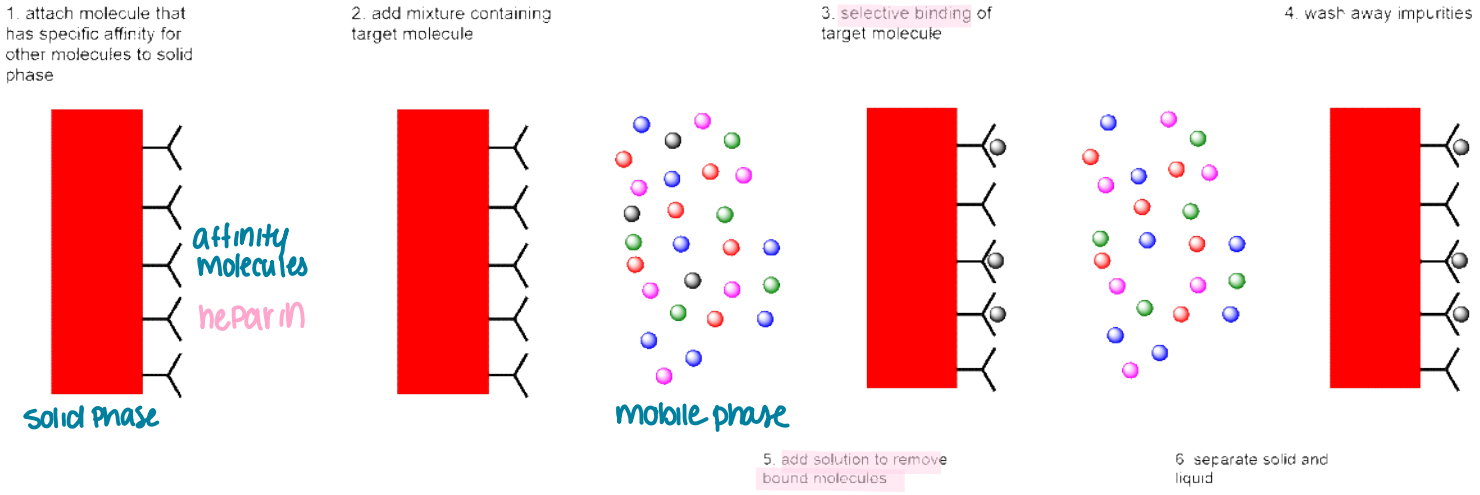
- virus is v hard to grow in culture
  - need liver cells to grow the virus
  - but liver cells are v hard to grow in culture
  - so its impractical to use viral culture to make a vaccine
- in the 60s, it was known that the surface antigen is immunogenic
  - could HBsAg become a vaccine? vaccinate using protein only
    - the protein is the part of the virus that the immune system responds to
  - could be a very safe way to vaccinate, impossible to get an infection from the vaccine
  - before this, all viral vaccines used entire virus
    - attenuated; non-virulent virus
    - dead virus; virus is intact but non-functional
    - split virus; killed and structure disrupted

- its impossible to manufacture the protein chemically bc it's a large protein
- impossible to grow virus in culture; cant grow liver cells in vitro
- in the 60s, the protein must be isolated from infected humans
    - antigen is paonly found in blood
    - infected people carry lots of the antigen in their blood
- first generation vaccine (1981)
    - collect plasma from infected people (IV drug users and homosexual men)
    - isolated viral protein aggregates from plasma
    - purified the protein aggregates
    - protein aggregates used to vaccinate
    - purification problem: the proteins are designed to reside in membrane
        - must retain small amounts of lipid to maintain proper protein structure and be immunogenic
- vaccine purification:
    - HBsAg requires lipids to remain immunogenic
        - protein is embedded in artificial membrane
        - cant purify 100%; purifying a mixture, not a single chemical substance
    - plasma dialysis from infected person to separate liquids from solids
    - affinity chromatography; sulphated heparin
        - a way to purify antigen and keep it with bits of lipid
    - HBsAg precipitated using the solvent PEG 6000 (polyethylene glycol mw 6000)
        - had substance dissolved in water
        - precipitated antigen by changing polarity of water using PEG 6000
    - HBsAg redissolved; used as vaccine
    - the final product contain proteins, lipids (serum)
- challenges in manufacture:
    - supply is limited
        - only small % of pop gives blood in NA
        - & blood must come from a carrier (chronic infection); this is <1% of pop
    - bc of limited supply, price of vaccine was 50X higher than other vaccines
        - usually $2/dose, hep B was $100/dose
- limited access to vaccine in 1981-1983
    - small amounts available bc limited source and $$
    - only given to individuals at moderate or high risk for hep B
        - health care professionals
        - infants born to hep B positive mothers
        - people with history of STDs & IV drug users
- emergence of new virus (HIV)
    - first AIDS case diagnosed in 1981; immune collapse, kaposi's sarcoma (rare skin cancer)
    - 100% fatal
    - spread by contact with body fluids
        - unsafe sexual practice, IV drug use, congenital infection of newborns
- there was a concern for the safety of the hep B vaccine bc blood was sourced from homosexual men and IV drug users, the exact groups that were at greater risk for developing AIDS
- AIDS/HIV changed blood donation and use
    - before 1983; show up and donate blood
        - blood was screened after collection, but limited number of tests available
    - after 1983; donors undergo pre-screening
        - certain groups excluded from donation
        - blood also screened after collection, larger number of tests available
        - ex of screening question: have you spent time in the UK during mad cow disease outbreak?

- before 70s, blood donors were paid in the US (where blood was collected for hep B vaccine)
  - donors were often people who needed money (homeless, IV drug users)
    - higher percentage of blood born diseases in these groups
  - volunteer system introduced in 1970
  - FDA (1978) required blood to be labelled as paid or volunteer
- screening of bloody supply in the US:
  - before 1970: blood type only
  - 1971: test for hep B available, but not applied to all blood collected
  - 1972: FDA begins regulating blood donation, nationalized standards
  - 1983: first warning about AIDS from blood transfusion; removed hep B vaccine
  - 1985: HIV test introduced (antibody test)
  - 2002: first use of PCR to test (HIV, hep C)
- blood derived products (treatments):
  - whole blood, plasma, platelets, proteins, hormones, hep B vaccine
  - after 1983, companies & hospital became very discriminating about use of blood products
- hep B vaccine was no longer used
  - donors for hepatitis proteins were in high risk groups for AIDS
  - purification doesn't remove all contaminants; bc vaccine was a large protein
    - risk of contamination with large biomolecules or viruses (today viral risk is negligible)
- need for an alternative
- second generation vaccine:
  - produce antigen protein using recombinant methods
  - no risk of viral infections
  - large quantities available
- preparation of yeast transforming vector:
- vaccine processing: (don't memorize sequence)
  - yeast cells are disrupted by sonication
  - ion exchange
  - affinity chromatography
  - gel filtration
  - $Al(OH)_3$ added (adjuvant)
    - adjuvants reduce # of booster shots
    - give stronger immune response
  - preservative may be added
- affinity chromatography
  - use stationary phase that has specific affinity for your molecule or portion of the molecule that you're trying to purify
    - stationary phase is coated with antibodies
    - or use a tight-binding molecule; common one is biotin-streptavidin
      - streptavidin has a very tight binding constant for biotin
  - for HBsAg, heparin-sulfate is used
    - heparin is complex carbohydrate found in blood



- affinity purification:



- recombivax (1986); the first recombinant vaccine
  - no risk of blood-borne infection
    - produced in yeast cells
  - no risk of hep B infection
    - no intact virus present
  - no supply issues
    - yeast is easy to grow in required amounts
- initial benefits of the vaccine were limited
  - only a small % of population is at risk for hep B
  - indicated for those at moderate to high risk
  - health care professionals (low-mod risk)
  - infants born to hep B positive mothers (mod risk)
  - ppl with history of STDs (high risk)
  - IV drug users (high risk)
  - highest risk groups had the lowest vaccination rates
    - vaccination rates were high for medical staff and infants
    - vaccination rates were low for people with history of STDs and IV drug users
- vaccine was underused
  - had a way to supply a lot, but underused
  - FDA was targeting high risk groups, but they were the least likely to get vaccinated
  - the vaccine also required 3 injections, this is hard to do in high risk groups
- so the CDC recommended all children get vaccinated in 1990
  - its hard to get adults in high risk groups to vaccinate
  - infected children tend to get chronic infections
    - become carriers
    - capable of infecting other for a lifetime
    - may or may not engage in high risk behaviour later in life
  - vaccinate all to achieve herd immunity
- vaccine required 3 injections; first given to newborns before leaving the hospital (partial protection) and the other two injection administered over 2-3 years
- twinrix vaccine has hep A vaccine (killed virus) and a hep B vaccine (recombinant protein antigen)
  - recommended to get this before you travel to developing countries
  - requires two injections

- SUMMARY: types of vaccines available
  - attenuate virus or bacteria
    - small risk of infection if immune system compromised
    - difficult to develop for high risk organisms (ebola)
      - even when dead, can have proteins on them that are problematic
  - killed virus or bacteria
    - extremely small risk of infection; reduced by splitting
    - manufacture defects
    - difficult for high risk organisms (ebola)
  - recombinant protein from virus or bacteria
    - no risk of infection
    - safe for high risk organisms (ebola, HIV)
- SUMMARY: recombinant vaccines available today
  - hepatitis virus; liver disease and liver cancer
    - twinrix
  - papilloma virus; cervical cancer
  - varicella zoster virus; singles
  - influenza (flu) virus; a specialized one avoids egg allergy
  - pneumococcal bacteria (23 species); various respiratory infections
    - prevnar
  - meningococcal bacteria (4 species); meningitis
  - bordetella pertussis; whooping cough