Lecture 12

- For eukaryotes and prokaryotes the major control point of gene regulation is transcriptional initiation
  - ~95% of cells are regulated at the stage of initiation
- Prokaryotes --> RNA polymerase directly contacts the promoter
- Eukaryotes --> there are many ancillary proteins that aid in transcription that recognize many different sequences that are distinct from the actual part that is bound by the polymerase and these factors recruit the polymerase
  - regions necessary for transcription are orders of magnitude larger than needed for prokaryotic transcription.
- Prokaryotic
  - region is not defined, usually found by taking segments upstream and nibbling away parts of the upstream region with nucleases and seeing what is the shortest piece that can support transcription
  - what does the polymerase recognize?
    - Not all 60 base pairs have to be defined to be a part of the promoter
    - the polymerase has a DNA binding domain, which is usually contributed by the β and β' subunits of the polymerase and not every nucleotide is necessary for these contacts. The actual sequences that get recognized, in fact, are two hexameric sequences that are separated by a fixed number of base pairs between them to form to the shape of the polymerase
  - one specific base used for starting and 4.2 million - 1 bases that are wrong start
    - for any given promoter your looking for a given promoter
    - cannot design a protein that is infinitely long to contact a DNA sequence that is infinitely long so that there would be infinite energies to bind a specific sequence and no energy to bind the non-specific sequence
    - in fact, even for proteins that bind DNA specifically, a lot of the contacts involve the phosphodiester backbone, which is shared by all DNA sequences
      - there is always an affinity for non-specific DNA
      - Normally there is orders of magnitude higher specificity for the correct sequence, but that is counter balanced by the amount non-specific sequence
        - at any given time the regulatory molecules are usually found at the wrong sites.
        - Proteins finding their specific sequences
          - trial and error (very slow and not probably method)

        - binds non-specifically and some how "skids" along the DNA looking for the correct sequence (no evidence of this)
        - start out making contact at non-specific sequences, while holding onto these sequences they do a "2d walk" along the genome and exchange a specific sequence for another sequence for the non-specific sequence (this results in less volume to explore and makes this model most probable)
            - Often proteins have different conformation or require auxiliary cadherins that make them have a higher affinity for their specific site.
- Proteins are not infinitely long because if they were it would take longer to find their complementary strand and to make the correct conformational shape.
- protein size is consistent, the difference between eukaryotes is their amount of subunits.
- Which is a trans-acting component to transcription? Is it the site that the polymerase recognizes or the site itself?
    - trans acting factor - acts on the opposite site
    - cis acting component - on the DNA
    - trans acting factors are diffusable so that organisms that are diploid or that you makes things, that you can have a protein that is encoded by one chromosome act on the opposite chromosome because it is diffusable.
        - proteins and regulatory RNAs are examples of transacting factors
        - Any DNA binding site or RNA binding has to act where it exist, the binding site or promoter doesn't diffuse away. The promoter is always attached to the gene so it is cis acting
            - binding sites are cis acting components, and the factors that recognize them are trans-acting components.
- Fig 11.16 Bacterial RNA polymerase consists of multisubunit
    - bacterial RNA core polymerase are ~500 kd multisubunit complexes with the general structure $\alpha 2\beta\beta'$
- Polymerase is composed of $\alpha 2\beta\beta'$, 2 subunits α that assemble the polymerase and β and β' make contact with the DNA. θ is the catalytic component, it is part of the enzyme that adds ribonucleotides to the RNA
    - σ - only used at initiation, recognizes the promoter
        - changes the binding properties of polymerase at initiation, and making the polymerase different at elongation
        - component that recognizes the promoter
- polymerase does not move off of the promoter for the first 9 nucleotides

- how does it keep making more nucleotides if it is not advancing?
  - DNA gets "scrunched" into the promoter, as it gets scrunched the polymerase gets shorter and shorter to accommodate the DNA that is being scrunched within it so that it can synthesize new bases opposite a different template base. DNA is getting more and more scrunched, by the time it is ready to elongate the polymerase has shrunk in size.
- After the first dinucleotides, the σ either changes conformation so that it lets go of the promoter where it falls off entirely, and if you added 9 nucleotides and made it to the 10th nucleotide then you have entered the elongation phase. Now the binding affinity has changed dramatically
- translocation - making contacts with a certain sequence on the template, the DNA is going to advance relative the DNA template. Breaking contacts and creating new ones as the nucleotide entry site is moving on the polymerase to the next part of the template.
  - When you break the contacts, if there was nothing holding on then the RNA would fall off every nucleotide, you want the RNA to stay attached to the polymerase
    - protein bridge is a component of the polymerase, which is apart of the β subunit that changes conformation so that while the sequence is translocating it is still making contact to the last nucleotide added so that the RNA doesn't fall off while this process is occurring
  - Fig 11.15 A protein bridge changes conformation to control the energy of nucleotides to the active sites
  - fig 11.17 DNA is bound in a channel and is contracted by both β and β' subunit
  - RNA polymerase consists of the core enzyme and σ factor
    - bacterial RNA polymerase can be divided into the α2ββ' core enzyme
  - β and β' subunit put the polymerase on the DNA
  - polymerase proper melts the DNA
  - α2ββ' core that is needed to continue elongation
    - σ is only used for initiation
    - when you have a completed α2ββ'σ (Holoenzyme)
      - the holoenzyme is the only structure of the polymerase that can recognize promoters
  - closed complex - DNA remains double stranded
  - open complex - DNA is melted in a short region
    - found when your starting to do initiation
  - loose binding site - core polymerase bound to nonspecific DNA sequence
    - called this because you don't want the core during elongation to dissociate from the template until it hits a termination signal sequence. Want it to get of the promoter,