Lecture #7
The Cell as a Machine

RNA Transcription

Taken from Alberts et al., Chapters 6 & 7

Summary: How is the right DNA read and transcribed into protein? With the definition of the expressed transcription factors and better data on the changes in transcription patterns, we can hope to understand what primary signals govern expression. Binding of several factors seems essential to have sufficient specificity and regulatory factors.

Problems:

A) Identify regions of the genome to be transcribed in general.
B) Activate transcription and processing at the correct time.
C) Inactivate transcription of those mRNAs where there is enough protein
D) Coordinate expression with other events in the cell cycle.
E) Ability to change phenotype
F) Control synthesis of other RNAs
G) Respond to feedback controls

Before we discuss the actual machinery that transcribes the DNA code into the RNA copy of that information, we will consider the general question of protein production in cells. The reason for transcribing DNA into RNA is to ultimately produce functional protein complexes. As in a well run automobile production plant, to continue our automobile theme, the cell needs to make the necessary parts (proteins) for the machines as it needs those machines. Production of an excess of one part over the others that make a given machine is wasteful and a shortage of an essential part will cause the machine to not work. Further, you need to have a quality control system that checks every part for functionality as it is made. In most eukaryotes, there are a number of different machines that need to be produced. At the most elemental level, the cell cycle has machines that are needed for mitosis only, others that are active in the DNA synthesis phase and still others that are specific for cytokinesis. In a multicellular organism, there are specialized machines that differentiate one cell from another and those should be produced in the correct cell but not in the incorrect cell.

Because it takes a lot of energy to produce a protein (at least 2 ATP molecules on average are needed per amino acid in a mature protein), there is a real premium placed on the efficient production of proteins. Thus, the cell carefully decides which protein to produce and how much of that protein to produce. At the level of RNA production, this means that the genes being transcribed are coordinately regulated.

RNA transcription factors (specificity 2 sites per nucleus or 30 picomolar)

Upstream sequences of the genes need to be scanned by activation complexes for the presence of sequences that bind factors involved in activating transcription. As it turns out, there are many transcriptional factors that stay bound to the DNA even when the message is not being transcribed, which obviates the need to scan for those regions of

DNA. The question of activation then changes from one of finding the sequence to finding the specific transcription factor, which would be easier.

RNA polymerase II is the polymerase that synthesizes mRNAs and it requires several activation factors before it can start the transcription process. Once a region of DNA is activated, one polymerase after another starts at the same site and moves through the gene, creating a Christmas tree-like array of progressively longer transcripts of the gene as you move away from the origination site. Error rate is 1 in 10^4 bases. 10% of genome is for RNA that is not translated into protein. 3-5% of total RNA is mRNA and most is present as rRNA.

Recent studies are focusing on the real time observation of the transcription and RNA processing events using GFP tags (Tsukamoto et al., 2000) (Kues et al., 2001). These studies reveal that the complexes are very dynamic and that soluble species can diffuse rapidly within the nucleus.

Questions:

1. For every 200 bp of double stranded DNA, there is one nucleosome and nucleosomes bind with an effective binding constant greater than 10^9 M^{-1}.($\Delta G = -RT \ln Ka$, where R = 8.31 J/°K mole) Suppose that you want to screen all of the DNA of an average chromosome of 150 million base pairs. How many ATP molecules will be hydrolyzed in the screen, assuming that the energy of ATP hydrolysis is 29 kJ/mole. (see the notes for the treatment of the energy of binding)
2. If we have an mRNA that is needed at the level of ten copies per cell on average and the half-life of the mRNA in cytoplasm is ten minutes, then what fraction of the time should the gene be activated. Assume that the polymerase makes RNA at a rate of 20 b/sec, that the original message is 6000 bases long, and that the next polymerase can start after the previous polymerase has moved 500 bp.

Kues, T., A. Dickmanns, R. Luhrmann, R. Peters, and U. Kubitscheck. 2001. High intranuclear mobility and dynamic clustering of the splicing factor U1 snRNP observed by single particle tracking.PG - 12021-6. *Proc Natl Acad Sci U S A*. 98.

Tsukamoto, T., N. Hashiguchi, S.M. Janicki, T. Tumbar, A.S. Belmont, and D.L. Spector. 2000. Visualization of gene activity in living cells. *Nat Cell Biol.* 2:871-8.