7 Transcriptional Control of Gene Expression

Review the Concepts

1. In glucose media, *lac* repressor bound to the *lac* operator sterically blocks initiation by RNA polymerase: *lac* operon expression is repressed. After shifting to lactose media, lactose is transported into the cell where it binds the *lac* repressor. This causes the repressor to dissociate from the *lac* operator, allowing RNA polymerase to bind to the *lac* promoter and initiate transcription. In addition, the cell synthesizes cAMP because of the low concentration of glucose. cAMP binds to the catabolite activator protein (CAP), causing it to bind to the CAP site upstream from the *lac* promoter. The *lac* promoter is a weak promoter in the absence of CAP binding; CAP bound at the CAP site stimulates RNA polymerase binding and transcription initiation at the *lac* promoter, resulting in a high rate of transcription from the *lac* promoter and maximal induction of the *lac* operon.

2. In two-component regulatory systems, one protein acts as a sensor and the other protein is a response regulator. In *E. coli* the proteins PhoR (the sensor) and PhoB (the response regulator) regulate transcription of target genes in response to the free phosphate concentration. PhoR is a transmembrane protein, whose periplasmic domain binds phosphate and whose cytoplasmic domain has protein kinase activity. Under low phosphate conditions, PhoR undergoes a conformational change that activates the protein kinase activity and results in the transfer of a phosphate group to Pho B. The phosphorylated form of PhoB then induces transcription from several genes that help the cell respond to low phosphate conditions.

3. RNA polymerase I is responsible for transcribing 18S and 28S rRNA genes. RNA polymerase II is responsible for mRNA transcription. RNA polymerase III transcribes tRNAs, 5S rRNAs, and several other small RNAs. Since RNA polymerase II is uniquely inhibited by a low concentration of α-amanitin, one can determine if any gene requires this polymerase by measuring gene transcription in the presence and absence of this compound. If RNA polymerase II is responsible for transcribing a gene, then transcription should only occur in the absence of α-amanitin.

4. The CTD becomes phosphorylated by a subunit of TFIIH during transcriptional initiation. The CTD is then further phosphorylated by cyclin T-CDK9. CTD phosphates are removed when RNA polymerase II terminates transcription.

5. TATA boxes, initiators, and CpG islands are all promoter elements. The TATA box was the first to be identified because it is found in the promoters of most genes expressed at high level. These were the first genes to be subjected to in vitro transcription. Since the TATA box occurs at a fixed location relative to the transcription start site (~30 bases upstream), it was relatively easy to recognize in the DNA sequence.

6. To identify DNA-control elements within promoter regions, investigators utilize 5′ deletion mutants and linker scanning mutants. These mutants contain a loss of sequence in the promoter region. The wild type DNA sequence and mutated DNA are separately introduced into cultured cells, or transgenic animals are generated with the wild-type and mutant DNA. Then the expression level of the associated gene or a fused reporter gene is assayed. When an activator binding site is deleted, expression is reduced. When a repressor binding site is deleted, expression is increased.

7. Promoter-proximal elements are located within ~200 base pairs of the transcription start site. Enhancers are located at greater distances, either upstream or downstream of the transcription start site. Enhancers continue to activate transcription

when they are inverted or moved many kilobases from the transcription start site.

8. Once a putative control region is identified, DNA footprinting with a nuclear extract can identify the precise DNA sequence that is bound by a protein in the extract. This assay depends on the ability of protein factors to "protect" DNA from DNaseI digestion. The electrophoretic mobility shift assay can also be used to determine whether proteins in a cell extract bind specifically to a DNA sequence within the probe used. In this assay, an extract is incubated with a labeled fragment of DNA. If protein binds to the DNA, then it "shifts" in terms of its migration on a polyacrylamide gel. This technique is often used as an assay to purify DNA-binding proteins.

9. Transcriptional activators and repressors contain a modular structure in which one or more transcriptional activation or repression domains are connected to a sequence-specific DNA-binding domain, usually through a flexible domain.

10. In patients with Wilms's tumor both copies of the WT1 gene are nonfunctional in tumor cells, so no WT1 transcriptional repressor is made. Since this repressor normally binds to the control region of the *EGR-1* gene and represses its transcription, Wilms's tumor cells have abnormally high levels of EGR-1. This high EGR-1 level results in aberrant transcription of growth control genes in kidney cells, contributing to the development of kidney tumors.

11. CREB binding to its co-activator (CBP) is regulated by camp, which stimulates phosphorylation of CREB. The phosphorylated acidic activation domain within CREB is a random coil in the absence of CBP. However, in the presence of CBP the phosphorylated activation domain undergoes a conformational change to form two α helices that wrap around a larger globular domain of the co-activator. In contrast, nuclear receptors contain a larger activation domain that is regulated by the binding of a hydrophobic ligand. When ligand binds to these domains, they undergo a conformational change that generates a groove in the globular activation domain that binds a short α helix in a co-activator. Thus, while the phospho-CREB activation is a relatively short random coil until it interacts with a larger globular domain in a co-activator, nuclear receptors have a large, folded, inactive activation domain that undergoes a conformational change allowing it to bind a short α helix in a co-activator.

12. The first protein to bind to a RNA polymerase II promoter is the TATA box–binding protein (TBP), a subunit of TFIID. This protein folds into a saddle-like structure which binds to the minor groove of DNA near the TATA box and bends the DNA. TFIIB then binds and makes contact with both TBP and DNA on either side of the TATA box. TFIIF and Pol II bind and Pol II is positioned over the start site. TFIIE and TFIIH bind and a helicase activity unwinds the DNA generating an "open" complex with the template strand in the active site of the polymerase.

13. Integration of gene *X* near the telomere is not ideal for good expression of gene *X*. Telomeres are usually contained in heterochromatin, which is tightly packed and less accessible for the transcriptional machinery. If the yeast line used for expression contained mutations in the H3 and H4 histone genes, the outcome could be different, depending on the specific mutations. For example, if the DNA sequence encoding lysines in the histone N-termini were mutated so that glycine residues were substituted in their place, then repression of gene *X* would not take place. Repression would not take place in such a mutant because the glycine residues are not positively charged, similar to acetylated lysine residues. This prevents binding by SIR3 and SIR4, preventing the formation of heterochromatin. The resulting "open" chromatin structure at the telomere would facilitate RNA polymerase II and general transcription factor binding, allowing gene expression.

14. A good prediction is that STICKY functions as a transcriptional repressor. Repressors contain two domains, one that binds DNA and a second that represses transcription. The bHLH domain is a DNA-binding domain that has been found in many different transcription factors. The Sin3-interacting domain is likely to associate with a

Sin3-containing histone deacetylase complex. This complex can repress transcription because deacetylation of histones promotes a more closed chromatin conformation.

15. A two-hybrid assay relies on unique yeast vectors, one referred as the "bait" and the other the "fish." The bait vector contains sequence encoding a DNA-binding domain (BD) and flexible linker region followed by a multiple cloning site (MCS). A cDNA encoding a known protein or protein domain is inserted into the MCS, in the proper frame to maintain the correct coding sequence. Four bait vectors need to be created for our experiment, one encoding the full-length GR, and three others, each encompassing one of the modular domains in the protein. The bait vector also expresses a wild-type tryptophan (*TRP*) gene, allowing *trp*$^-$ yeast cells harboring this plasmid to propagate in media lacking *TRP*. A cDNA library generated from the pituitary cells is cloned into the MCS of multiple copies of another vector, which encodes sequence for a flexible linker and a strong activation domain (AD), needed to recruit co-activators and the transcription pre-initiation complex. This fish vector also contains a wild-type leucine (*LEU*) gene, allowing *leu*$^-$ cells containing the plasmid to grow in media lacking *LEU*. To identify a protein-protein interaction, the bait vector encoding the full-length GR and a library of fish vectors are transfected into engineered yeast cells containing a gene required for histidine (*HIS*) synthesis under the control of a UAS with binding sites for the bait vector's DNA-binding domain. Cells are plated on media lacking leucine and tryptophan to maintain the bait and fish vectors in cells, and lacking histidine to prevent cells from propagating unless the *HIS* gene is transcribed. *HIS* transcription requires that the DNA-binding domain of the bait hybrid bind the UAS and the GR portion of the bait fusion protein has interacted with a fish protein fused in frame to the activation domain. Cells expressing both bait and fish plasmids survive and grow on media lacking *HIS* only if an interaction has occurred between GR and an interacting fusion fish protein. Recovery of the fish vector encoding the interacting protein and its subsequent sequencing reveals that Hsp90 interacts with the GR. Having used the full-length GR to identify Hsp90, the experiment would be repeated with the three GR domain-encoding bait vectors. Instead, however, of using the entire cDNA library, each bait vector would be cotransfected into cells with the fish plasmid encoding Hsp90. Results would show that the GR ligand-binding domain is the one that binds Hsp90.

16. The heat-shock genes are induced when cells are exposed to harsh conditions, such as high temperature, that cause the denaturation of proteins. In a cell growing at normal temperature an RNA polymerase II molecule has initiated transcription of the Hsp70 gene, but paused after the transcription of ~25 nucleotides. When the cell is exposed to high temperature, the heat-shock transcription factor is converted to an active comformation, binds to the Hsp70 promoter-proximal region, and stimulates the paused polymerase to continue chain elongation and transcribe the full length of the Hsp70 gene. This mechanism shortens the time for RNA expression because no time is required to assemble a transcription pre-initiation complex.

Analyze the Data

a. In untreated cells (t = 0 minutes), both wild type and CARA cells transcribed equivalent amounts of ribosomal RNAs, suggesting that the mutant Pol I functions normally under control conditions. This finding is consistent with the observation that the growth rate of yeast harboring this mutation was indistinguishable from wild type cells. However, when the drug rapamycin was added, Pol I transcription in wild-type cells was completely inhibited by 100 minutes after addition of the drug. In contrast, in the CARA cells, Pol I transcription was much less inhibited and continued at significant level even at 100 minutes after addition of rapamycin. Consequently, Pol I transcription in the CARA cells was only partially inhibited by rapamycin.

b. As for rRNA precursor synthesis by Pol I analyzed in (a), both ribosomal protein mRNAs transcribed by Pol II and 5S rRNA synthesis by Pol III were rapidly inhibited by the addition of rapamycin in wild type cells. Also, as observed for Pol I in (a), transcription of ribosomal protein and 5S rRNA

genes was partially resistant to rapamycin and continued for a longer period after the addition of rapamycin in CARA cells compared to wild-type cells. Since the genetic modification of CARA cells directly affects Pol I, but not Pol II or Pol III, these differences from wild-type cells are probably a consequence of the continued Pol I transcription of the large rRNA precursor in CARA cells. Thus these results indicate that transcription of ribosomal protein mRNAs by Pol II and of 5S rRNA by Pol III is coordinated with the supply of the large rRNA precursor. However, the molecular mechanisms that coordinate the transcription of these genes by Pol II and Pol III with the supply of the large rRNA precursor transcribed by Pol I remain to be discovered. This strategy for regulating the transcription of ribosomal protein genes and 5S rRNA genes seems logical. Since the ribosomal proteins and 5S rRNA are assembled onto the large rRNAs transcribed by Pol I during the synthesis of the ribosomal subunits, it makes sense for the cell to increase synthesis these proteins and 5S rRNA in response to the supply of the large rRNAs.

c. In rich media, the CARA cells express the majority of genes, including the ribosomal protein genes, at a level similar to that of wild-type cells. Only a very small number of genes were expressed at more than twofold or less than one-half the level in wild type cells. However, when the cells were shifted to poor media, the ribosomal mRNAs, but not other mRNAs, were expressed at a higher (twofold) level in the CARA cells than in the wild-type cells. Thus these mRNAs were over-represented in CARA cells under conditions that do not involve the use of drugs, but simply a decrease in nutrient supply. This over-representation in the mutant cells may be explained as follows: In wild-type cells, the shift to the poor media results in repression of Pol I synthesis of the large ribosomal RNAs and, coordinately, repression of Pol II synthesis of mRNAs encoding ribosomal proteins. In CARA cells, where the genetic manipulation affects Pol I and not Pol II, the continued expression of Pol I in low nutrient medium resulted in increased transcription of the ribosomal protein genes, confirming that transcription of ribosomal protein genes is stimulated in response to the production of the large rRNAs by Pol I.