Eukaryotic transcriptional control: major considerations

1. Interplay among multiple general transcription factors; activators/repressors; mediator/coactivators
2. Multiple regulatory sequences: proximal and distant
3. Chromatin and its impact on transcription
4. Co-transcriptional RNA processing
5. Regulation of transcriptional regulators

Sequence-specific transcription factors are modular

Modular structure of Sp1
Experiments to map the DNA-binding and activation domain of yeast GAL4 protein

<table>
<thead>
<tr>
<th>Wild-type and mutant GAL4 proteins</th>
<th>Binding to UAS\text{gal}</th>
<th>β-galactosidase activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild-type N-terminal deletion</td>
<td>74</td>
<td>738 823</td>
</tr>
<tr>
<td></td>
<td>74</td>
<td>881</td>
</tr>
<tr>
<td>N- and C-terminal deletion mutants</td>
<td>74</td>
<td>848</td>
</tr>
<tr>
<td></td>
<td>74</td>
<td>823</td>
</tr>
<tr>
<td></td>
<td>74</td>
<td>792</td>
</tr>
<tr>
<td></td>
<td>74</td>
<td>755</td>
</tr>
<tr>
<td></td>
<td>74</td>
<td>725</td>
</tr>
<tr>
<td></td>
<td>74</td>
<td>692</td>
</tr>
<tr>
<td>Internal deletion mutants</td>
<td>74</td>
<td>684</td>
</tr>
<tr>
<td></td>
<td>74</td>
<td>738 881</td>
</tr>
<tr>
<td></td>
<td>74</td>
<td>681</td>
</tr>
<tr>
<td></td>
<td>74</td>
<td>760 881</td>
</tr>
</tbody>
</table>

DNA-binding domains can be classified into many structural types
One type of zinc finger protein (C2H2)

This protein belongs to the Cys-Cys-His-His family of zinc finger proteins, named after the amino acids that grasp the zinc. This zinc finger is from a frog protein of unknown function. (A) Schematic drawing of the amino acid sequence of the zinc finger. (B) The three-dimensional structure of the zinc finger is constructed from an antiparallel \( \beta \) sheet (amino acids 1 to 10) followed by an \( \alpha \) helix (amino acids 12 to 24). The four amino acids that bind the zinc (Cys 3, Cys 6, His 19, and His 23) hold one end of the \( \alpha \) helix firmly to one end of the \( \beta \) sheet. (Adapted from M.S. Lee et al., *Science* 245:635-637, 1989. © 1989 the AAAS.)

The binding of transcription factors to the major groove of DNA

- As with most bacterial activators and repressors, \( \alpha \) helices in the DNA-binding domain of eukaryotic transcription factors are often oriented so that they lie in the major groove of DNA helix where atoms of protein and DNA make contact through specific H-bonds and van der Waals interactions.
- Typically, a protein-DNA interface consists of 10 to 20 such contacts, involving different amino acids, each contributing to the binding energy of the protein-DNA interaction.
DNA binding by a zinc finger protein

(A) The structure of a fragment of a mouse gene regulatory protein bound to a specific DNA site. This protein recognizes DNA using three zinc fingers of the Cys-Cys-His-His type arranged as direct repeats. (B) The three fingers have similar amino acid sequences and contact the DNA in similar ways. In both (A) and (B) the zinc atom in each finger is represented by a small sphere. (Adapted from N. Pavletich and C. Pabo, Science252:810-817, 1991. © 1991 the AAAS.)

The basic helix-loop-helix protein Max binds DNA as a dimer
Leucine zipper (aka b-Zip) proteins (e.g. Fos, Jun, & yeast GCN4) bind DNA as dimers

Leu residues at every seventh position down one side of the α-helix. Two α-helical monomers form a coiled-coil dimer. Basic amino acid residues N-terminal to the leucine zipper form the DNA-binding domain.

Heterodimeric transcription factors increase regulatory diversity and gene-control options

(a) Many transcription factors (e.g. b-Zip and helix-loop-helix proteins) can form both homodimers or heterodimers with other members of the same class.

(b) In the hypothetical example shown, transcription factors A, B, and C can each interact with each other, permitting the three factors to bind to six different DNA sequences (sites 1–6) and creating six combinations of activation domains. (Note that each binding site is divided into two half-sites, and that a single heterodimeric factor contains the activation domains of each of its constituent monomers.)

(c) When an inhibitory factor (green) is expressed that interacts only with factor A, binding to sites 1, 4, and 5 is inhibited, but binding to sites 2, 3, and 6 is unaffected.
Inhibitory regulation by truncated HLH proteins

The HLH motif is responsible for both dimerization and DNA binding. On the left, an HLH homodimer recognizes a symmetric DNA sequence. On the right, the binding of a full-length HLH protein to a truncated HLH protein that lacks the DNA-binding helix generates a heterodimer that is unable to bind DNA tightly. If present in excess, the truncated protein molecule blocks the homodimerization of the full-length HLH protein and thereby prevents it from binding to DNA.

True activation vs. antirepression

Gene activity

Activated State

Derepressed State (decondensed euchromatin)

Inactive Ground State (template DNA in the form of condensed chromatin in heterochromatin region)
Nucleosomes in condensed chromatin inhibit transcription at multiple stages

Workman and Kingston

How to activate chromatin for transcription

1. Covalently modify histone termini: e.g. H3 lysine9 acetylation

2. Move nucleosomes out of the way of the promoter in an ATP-dependent manner

3. Use histone modifications (histone code) to recruit other activator/co-activators
Acetylation induces a conformational change in the core histones

**EXAMPLE**

Lysine-rich tails bind tightly to DNA and repress transcription by blocking access of factors to the DNA template.

Acetylation of lysine residues alters chromatin structure and allows binding of transcription factors.

Note: acetylation neutralizes the positive charge of lysine

\[
\text{lysine}^+ \rightarrow \text{lysine}^\cdot
\]

\[
\text{HAT} \quad \text{CH}_3\text{C}--\text{CoA} \rightarrow \text{CoA}
\]

\[
\text{lysine}^\cdot \text{NH} \quad \text{CCH}_3
\]

\[
\text{O}
\]

**HAT: Histone Acetyltransferase**

Activator-directed hyperacetylation of histone N-terminal tails

Hyperacetylation of histones in the vicinity of the Gcn4-binding site facilitates access of general transcription factors required for initiation. Gcn5 is the catalytic subunit of the histone acetyltransferase (HAT) complex.
Repressor-directed deacetylation of histone N-terminal tails

Deacetylation of histone N-terminus on nucleosomes in the vicinity of the Ume6-binding site inhibits binding of general transcription factors at the TATA box, thereby repressing transcription. Rpd3 is the catalytic subunit of the histone deacetylase complex (HDAC).

Concerted Actions of Multiple Histone Modifying Enzymes in Gene Regulation

**Derepressed/open state**

A<sub>e</sub>

ARTKQTAR<sub>K</sub>STGGKAPRKLATKAK<sub>K</sub>SAP <sub>H3</sub>

histone lysine demethylase + HAT

histone deacetylase (HDAC) + histone methyltransferase (HMT)

**Inactive/condensed state**

A<sub>Me</sub>

ARTKQTAR<sub>K</sub>STGGKAPRKLATKAK<sub>K</sub>SAP <sub>H3</sub>
Chromatin-remodeling factors participate in activation at many promoters

- Many activation domains bind to chromatin-remodeling complexes, which serve as a type of co-activators, to stimulate transcription from chromatin templates.

- Chromatin-remodeling complexes in eukaryotes (ySwi/Snf; hSwi/Snf; hACF; RSF; etc) all contain a helicase/ATPase component to disrupt interactions between base-paired nucleic acids or between nucleic acids and proteins.

- The yeast Swi/Snf complex, the first well-characterized chromatin-remodeling complex, causes DNA bound to the surface of the histone octomer to transiently dissociate from the surface and translocate, allowing nucleosomes to “slide” along the DNA.

- Chromatin-remodeling complexes are required for many processes involving DNA.

ATP-dependent nucleosome “sliding” along DNA caused by chromatin-remodeling complexes

[Diagram showing ATP-dependent nucleosome sliding]
The mediator complex forms a molecular bridge between activation domains and Pol II

**S. cerevisiae mediator**

**Human mediator**

Control of eukaryotic transcription initiation: Ordered binding and function of activators and co-activators result in cooperative formation of a stable activated initiation complex

- Sequence-specific transcriptional activators; Chromatin remodelers; and Histone modifiers
- Several activators cooperatively interact with a single mediator complex
Regulation of cell-type-specific transcription by specific combinations of transcription factors

The TTR (transthyretin) gene is expressed only in hepatocytes but not in cells of the intestine and kidney, where AP1, HNF4 and C/EBP are expressed. All five activators are required to assemble an activated transcription initiation complex in a cooperative manner.

How do enhancers function in a gene-specific fashion?

A promoter and its enhancers are “cordoned off” from other promoter/enhancer elements by specialized Boundary or Insulator elements that are recognized by several nonhistone proteins (e.g. CTCF).
Boundary elements also prevent spreading of silenced and HP1-coated heterochromatin

The chromatin immunoprecipitation (ChIP) method
ChIP assay to distinguish promoter-bound versus elongating transcription proteins

Regulation of transcription at the stage of elongation

**HIV life cycle**

- Tat activates HIV-1 transcriptional elongation.
- Secreted by infected cells and taken up by uninfected bystander cells. Tat induces apoptosis in bystander cells.
Phosphorylation of Pol II CTD and negative elongation factors by P-TEFb stimulates transcriptional elongation

- HIV-1 transcription is exquisitely P-TEFb-dependent
- HIV-1 Tat & TAR recruit P-TEFb to the viral promoter
Regulation of Regulators

Examples:
- GATA-1
- CAP/nuclear hormone receptors
- NtrC/CREB
- Adenovirus E1A + CBP/p300
- NF-KB/glucocorticoid receptor

Cyclic AMP-Inducible Gene Expression—CREB Links cAMP Signals to Transcription

In animal cells, an elevation in the cytosolic cAMP level activates the transcription of specific target genes that contain a transcription factor (CREB)-binding site (CRE, cAMP responsive element). Regulation of gene expression by cAMP and CREB plays important roles in controlling cell proliferation as well as learning and memory.
Signal-induced degradation of a cytosolic inhibitor protein activates the NF-κB transcription factor

NF-κB, the master transcriptional regulator of the immune system (directly stimulates ~150 genes), is activated by inflammatory cytokines such as TNF-α and interleukin 1 (IL-1), which are released by nearby cells in response to infection. In addition to infection and inflammation, NF-κB can also be activated by other stressful situations, such as ionizing radiation.

**FIGURE 14-28.** NF-κB signaling pathway. In resting cells, the dimeric transcription factor NF-κB, composed of p60 and p65, is sequestered in the cytosol bound to the inhibitor IκB. Stimulation by TNFα or IL-1 induces activation of TAK1 kinase step 1, leading to activation of the trimeric IκB kinase step 2. Ionizing radiation and other stresses can directly activate IκB kinase by an unknown mechanism step 12. Following phosphorylation of IκB by IκB kinase and binding of E3 ubiquitin ligase step 3, polyubiquitination of IκB step 4 targets it for degradation by proteasomes step 5. The removal of IκB unmask the nuclear-localization signal (NLS) in both subunits of NFκB, allowing their translocation to the nucleus step 6. Here NFκB activates transcription of numerous target genes step 7, including the gene encoding the subunit of IκB, which acts to terminate signaling. See M. Kim and Y. Ben-Neriah, 2000, Annu. Rev. Immunol. 18:621, and R. Kuhn, F. Leufer, and B. Lemaire, 2001, Trends Immunol. 22:200.

Nuclear Receptor Superfamily: Transcription Factors Regulated by Lipid-Soluble Ligands Derived from Diet and Environment

Ligands: steroid and thyroid hormones; vitamins A and D; retinoids, etc.

**STEROID HORMONE**

- Cholesterol-derived hormones that have profound effects on gene transcription.
- Examples of steroid hormones are the glucocorticoids, such as cortisol; estrogens, such as β-estradiol; and androgens, such as testosterone.
- Cortisol became available shortly before the 1960 presidential election and may have had an important influence on the perceived outcome of the Kennedy-Nixon television debate. Kennedy suffered from Addison’s disease (inadequate adrenal function) at the time.
- Anabolic steroids, which are well known in athletics, help build muscle mass. They are related to the male sex hormone testosterone.
- Testicular feminization is a genetic condition (a mutation in the receptor for testosterone) in which genotypic (XY) males are unable to respond to testosterone and as a consequence develop the phenotypic characteristics of a female.
General design of transcription factors in nuclear-receptor superfamily

![Diagram of amino acid identity and domain structure for various nuclear receptors]

Experimental demonstration that hormone-binding domain of the glucocorticoid receptor (GR) mediates translocation to the nucleus in the presence of hormone

(a) 
![Immunofluorescence image without hormone](image1.png)

(b) 
![Immunofluorescence image with hormone](image2.png)

(c) 
![Immunofluorescence image control](image3.png)

Immunofluorescence: techniques that use antibodies chemically linked to a fluorescent dye to identify or quantify antigens in a tissue/cell sample.

Proteins expressed: N

- N-galactosidase
- Glucocorticoid receptor
- GR hormone-binding domain

<table>
<thead>
<tr>
<th>Protein</th>
<th>N</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>N-galactosidase</td>
<td>N</td>
<td>C</td>
</tr>
<tr>
<td>Glucocorticoid receptor</td>
<td>N</td>
<td>C</td>
</tr>
<tr>
<td>GR hormone-binding domain</td>
<td>N</td>
<td>C</td>
</tr>
</tbody>
</table>
**Future Perspectives**

- Discovery of new co-activators and co-repressors.
- Higher-order chromatin structure.
- Mechanism of integrating multiple signals.
- Cross talk with other nuclear processes.
- High throughput methods for studying gene expression
- Connections with human diseases.