I. *Cis*-acting sequences

core promoter: very near transcription start site, includes TATA box
promoter-proximal elements: 100-200 bp from start, includes CAAT box or GC-rich sequence

enhancers (for positive control) and **silencers** (for negative control) can act at large distances from the transcription start site, even downstream. The typically function even if they are moved or inverted, and will work to control expression off **heterologous** promoters.

A typical gene is regulated by complex combinations of enhancers and silencers.

Enhancers are thought to function by looping to interact with the core promoter.

II. *Trans*-acting factors (transcription factors)

Basal transcription factors (TFIIA, TFIIB, etc.) bind to the core promoter to assist RNA polymerase. TFIID includes **TBP** (TATA-binding protein) and **TAFs** (TBP-associated factors); TFIIH includes some DNA repair proteins.

Transcriptional activators = proteins that bind to enhancers to increase transcription at the associated promoter; often have the following domains:

DNA binding domain = binds to a specific nucleotide sequence in DNA;
activation domain = interacts with RNA polymerase to increase transcription; often acidic, asparagine-rich, or proline-rich;
dimerization domain = activators often work as homodimers or heterodimers;
ligand binding domain = some transcription factors are only functional when bound to an activating molecule (*e.g.*, steroid hormone receptors).

III. Common DNA binding domains

zinc finger – cysteine and histidine residues chelate a Zn^{++} atom.

example: steroid hormone receptors (testosterone and estrogen receptors); binding of hormone causes them to bind DNA to activate genes for sexual differentiation.

helix-loop-helix – two alpha helices separated by a short stretch of amino acids;

example: homeodomain proteins, which are important in specifying positional nformation during development (helix-loop-helix is a <u>structure</u>; homeodomain is one class of <u>amino acid sequence</u> that contains that structure). Genes that encode polypeptides with homeodomains are called homeobox (or Hox) genes.

IV. Mis-expression mutations

Chromosome rearrangements can place a structural gene adjacent to regulatory sequences of another gene, resulting in it being expressed at the wrong time or in the wrong place. These are often dominant gain-of-function mutations.

Antennapedia (*Antp*) is a homeobox gene that controls development of the abdominal segments in Drosophila – it is expressed in abdominal segments and is required to turn on all the genes required for formation of abdominal appendages. The $Antp^D$ mutation is an inversion that causes the gene to be expressed in the head. $Antp^D / Antp^+$ flies have legs instead of antennae (a homeotic transformation = one body part changed into another). $Antp^D / Antp^D$ flies die during embryonic development because they lack the normal expression of the gene in the abdomen. Thus, the mutation is dominant with respect to the antenna-to-leg transformation, but recessive with respect to embryonic viability.

Translocations between the human antibody genes (Cμ) on chromosome 14 and the *c-myc* transcription factor on chromosome 8 cause *c-myc* to be expressed in B cells. *c-myc* turns on genes involved in cell proliferation, Burkitt's causing lymphoma.

Controlled mis-expression is sometimes used to learn about gene function. Example: *Pax6* is thought to be required for eye development in mammals and in Drosophila. Mis-expression of the gene in Drosophila leads to "ectopic" eye tissue.

IV. Characterizing *cis* sequences by reporter gene expression

A method for mapping enhancers is to fuse the regulatory DNA for the gene being analyzed to a **reporter gene**. One then assays the product of the reporter gene. By fusing different segments of the regulatory DNA, one can map out which regions are required for which aspects of the expression pattern.

E. coli lacZ: beta-gal act on an artificial substrate called X-gal to generate a blue color
green fluorescent protein (GFP): a protein from jellyfish that fluoresces