Chapter 24: Promoters and Enhancers

- A typical gene transcribed by RNA polymerase II has a promoter that usually extends upstream from the site where transcription is initiated
  - the (+1) of transcription
- The promoter contains “several” short (<10 bp) sequence elements
  - bind transcription factors
  - dispersed over >200 bp
- An enhancer contains
  - An array short (<10 bp) sequence elements
  - also bind transcription factors
  - may be located several to many kb distant.
- Hypothesis regarding Enhancers
  - DNA may be coiled or otherwise rearranged so that transcription factors at the promoter and at the enhancer interact to form a large protein complex.

RNA Pol I:
- rRNA in the nucleolus

RNA Pol II:
- mRNA in the nucleoplasm

RNA Pol III:
- tRNAs and other small RNAs

RNA polymerase has >10 subunits

Pol I promotes have two sequence components

Transcription is controlled by a promoter and an enhancer

Enlarged view
There are three types of pol III promoters

Type 1
- Startpoint
- Box A
- Box C

Type 2
- Box A
- Box B

Type 3
- Oct
- PSE
- TATA

24.5 The Start point for RNA Polymerase II

- **A basal factor**:  
  - is a transcription factor required by RNA polymerase II
  - to form the initiation complex at all promoters.
  - Factors are identified as TFII X, where X is a letter.

- **The core promoter for RNA polymerase II**:  
  - is the minimal sequence at which the basal transcription apparatus can assemble.
  - It is typically ~40 bp long.
  - It typically includes a TATA box or an Inr and a DPE.
  - Some contain a TATA and an Inr (but no DPE)
  - Some promoters have none of these (TATA- and Inr-less promoters)
24.5 The Start point for RNA Polymerase II

- **The initiator (Inr):**
  - is the sequence of a pol II promoter between −3 and +5 and
  - has the general sequence YYANYYY. (C/T C/T A N A/T C/T C/T)
  - It is the simplest possible pol II promoter.

- **TATA box:**
  - is a conserved A·T-rich octamer found about 27 bp before the start point of each eukaryotic RNA polymerase II transcription unit;
  - It is involved in positioning the enzyme for correct initiation.

- **A TATA-less promoter:**
  - does not have a TATA box in the sequence upstream of its start point.

- **The DPE:**
  - Downstream Promoter Element
  - is a common component of RNA polymerase II promoters that do not contain a TATA box.
  - has the general sequence RGWYV(T) (A/G G A/T C/T A/C/G (T))

24.6 TBP Is a Component of TFII D and Binds the TATA Box

- **TFII D**
  - is the transcription factor that binds to the TATA sequence
  - ~27 bp upstream of the start point of promoters for RNA polymerase II.
  - It consists of TBP (TATA-binding protein)
  - and the TAF subunits that bind to TBP.

- **TAFs**
  - are the subunits of TFII D
  - assist TBP in binding to DNA.
  - They also provide points of contact for other components of the transcription apparatus.
24.6 TBP Is a Component of TF$_{II}$D and Binds the TATA Box

- TBP is a component of the positioning factor that is required for each type of RNA polymerase to bind its promoter.
- The factor for RNA polymerase II is TF$_{II}$D, which consists of TBP and 11 TAFs, with a total mass ~800 kD.
- TBP binds to the TATA box in the minor groove of DNA.
- TBP forms a saddle around the DNA and bends it by ~80°.
- Some of the TAFs resemble histones and may form a structure resembling a histone octamer.

24.7 The Basal Apparatus Assembles at the Promoter

- Binding of TF$_{II}$D to the promoter is the first step in initiation.
- Other transcription factors bind to the complex in a defined order, extending the length of the protected region on DNA.
- When RNA polymerase II binds to the complex, it initiates transcription.
24.9 Short Sequence Elements Bind Trans Activators

• An activator or Transactivator
  − is a protein that stimulates the expression of a gene;
  − typically by acting at a promoter to stimulate RNA polymerase.
  − In eukaryotes, the sequence to which it binds in the promoter is called a response element or a consensus sequence.
  − Examples include:

  • CAAT box
    − is part of a conserved sequence that may be located upstream of the start points of eukaryotic transcription units;
    − it is recognized by a large group of transcription factors.
    − CAAT enhancer binding proteins (C/EBP)

  • The GC box
    − is a common pol II promoter element
    − consisting of the consensus sequence GGGCGG.
    − Is bound by the Sp1 transactivator

Saturation mutagenesis of the upstream region of the β-globin promoter identifies three short regions (centered at -30, -75, and -90) that are needed to initiate transcription.

These correspond to the TATA, CAAT, and GC boxes.

In general, Promoters contain different combinations of TATA boxes, CAAT boxes, GC boxes, and other transcription factor binding sites.
24.11 Enhancers Contain the Same Elements That Are Found at Promoters

• An **enhancosome**
  - is a complex of transcription factors that assembles cooperatively at an enhancer.

• **Key Concepts**
  - Enhancers are made of the same short sequence elements that are found in promoters.
    - Transcription factor binding sites
  - The density of sequence components is greater in the enhancer than in the promoter
    - usually
  - Enhancers can be found
    - upstream (5′) or downstream (3′) of a promoter

• An enhancer contains several transcription factor binding sites

• The histogram plots the effect of all mutations that reduce enhancer function to <75% of wild type.

• Binding sites for proteins are indicated below the sequence.

---

**Enhancer action is independent of location**

- Upstream enhancer activates promoter
- Downstream enhancer activates promoter

**Enhancer activity requires proximity to the promoter**

- Enhancers usually work only in *cis* configuration with a target promoter.

- The principle is that an enhancer works in any situation in which it is constrained to be in proximity with the promoter.

- Enhancers Work by Increasing the Concentration of Activators Near the Promoter
24.13 CpG Islands Are Regulatory Targets

- **A CpG island**
  - is a stretch of 1–2 kb in a mammalian genome
  - that is rich in unmethylated CpG doublets.
- Demethylation at the 5’ end of the gene and the promoter region is necessary for transcription.
- CpG islands surround the promoters of constitutively expressed genes where they are unmethylated.
- They are also found at the promoters of some tissue-regulated genes.
- There are ~29,000 CpG islands in the human genome.
- Methylation of a CpG island usually prevents activation of a promoter within it.
  - Repression is caused by proteins that bind to methylated CpG doublets.
  - Or the methylation can block transcription factor binding
- Many cancers have aberrantly methylated CpG islands

CpG islands have clusters of CpG doublets

- The typical density of CpG doublets in mammalian DNA is ~1/100 bp,
  - as seen for a γ-globin gene.
- In a CpG-rich island,
  - the density is increased to >10 doublets/100 bp.
- The island in the APRT gene starts
  - ~100 bp upstream of the promoter
  - extends ~400 bp into the gene.
  - Causes these genes to be shut off inappropriately
  - Tumor Suppressor genes