Eukaryotic & Prokaryotic Transcription

RNA polymerases
RNA Polymerases

A. E. coli RNA polymerase

1. core enzyme = bb'(a)2 has catalytic activity but cannot recognize start site of transcription
   ~500,000 daltons
   dimensions: 100 X 100 X 160 angstroms
   requires Mg2+ for activity
   b' binds 2 Zn atoms

2. holoenzyme = core enzyme + sigma factor (s)
   carries out four functions:
   (i) template binding
   (ii) RNA chain initiation
   (iii) chain elongation
   (iv) chain termination
RNA Polymerases

B. Eukaryotic RNA polymerases (RNAP)
   1. 3 nuclear RNA polymerases
      a. RNAP I - transcribes rRNA genes
      b. RNAP II - transcribes mRNA genes
      c. RNAP III - transcribes tRNA, 5S rRNA, and other small RNA genes
      d. have 10-17 different subunits, large multisubunit complexes are functionally similar to E. coli RNA polymerase
      e. cannot bind to their respective promoters alone, but requires transcription factors for promoter specific recruitment
<table>
<thead>
<tr>
<th>S. cerevisiae</th>
<th>S. cerevisiae</th>
<th>S. cerevisiae</th>
<th>E. coli</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>RNAP I (14 subunits)</td>
<td>RNAP II (12 subunits)</td>
<td>RNAP III (15 subunits)</td>
<td>RNAP Core (5 subunits)</td>
<td></td>
</tr>
<tr>
<td>Rpa1 (A190)</td>
<td>Rbp1 (B220)</td>
<td>Rpc1 (C160)</td>
<td>β'</td>
<td>Core</td>
</tr>
<tr>
<td>Rpa2 (A135)</td>
<td>Rbp2 (B150)</td>
<td>Rpc2 (C128)</td>
<td>β</td>
<td>Core</td>
</tr>
<tr>
<td>Rpc5 (AC40)</td>
<td>Rpb3 (B44.5)</td>
<td>Rpc5 (AC40)</td>
<td>α</td>
<td>Core</td>
</tr>
<tr>
<td>Rpc9 (AC19)</td>
<td>Rpb11 (B13.6)</td>
<td>Rpc9 (AC19)</td>
<td>α</td>
<td>Core</td>
</tr>
<tr>
<td>Rbp6 (ABC23)</td>
<td>Rbp6 (ABC23)</td>
<td>Rpb6 (ABC23)</td>
<td>ω</td>
<td>Core/common</td>
</tr>
<tr>
<td>Rpb5 (ABC27)</td>
<td>Rpb5 (ABC27)</td>
<td>Rpb5 (ABC27)</td>
<td></td>
<td>Common</td>
</tr>
<tr>
<td>Rpb8 (ABC14.4)</td>
<td>Rpb8 (ABC14.4)</td>
<td>Rpb8 (ABC14.4)</td>
<td></td>
<td>Common</td>
</tr>
<tr>
<td>Rbp10 (ABC10β)</td>
<td>Rpb10 (ABC10β)</td>
<td>Rpb10 (ABC10β)</td>
<td></td>
<td>Common</td>
</tr>
<tr>
<td>Rbp12 (ABC10α)</td>
<td>Rpb12 (ABC10α)</td>
<td>Rpb12 (ABC10α)</td>
<td></td>
<td>Common</td>
</tr>
<tr>
<td>Rpa9 (A12.2)</td>
<td>Rpb9 (B12.6)</td>
<td>Rpc12 (C11)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rpa8 (A14)</td>
<td>Rpb4 (B32)</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rpa4 (A43)</td>
<td>Rpb7 (B16)</td>
<td>Rpc11 (C25)</td>
<td>+4 others</td>
<td></td>
</tr>
<tr>
<td>+2 others</td>
<td>+4 others</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*aHomologous subunits occupy the same row. In the alternative subunit names in parentheses, the letter(s) indicates the RNAPs in which the subunit is a component (A, B, and C for RNAPs I, II, and III) and the numbers indicate its approximate molecular mass in kD.

bCore: sequence partially homologous in all RNAPs; common: shared by all eukaryotic RNAPs.

cPotential homologs of Rbp4 and Rbp7.

dRpa3 (A49) and Rpa5 (A34.5) in RNAP I and Rpc3 (C74), Rpc4 (C53), Rpc6 (C34), and Rpc8 (C31) in RNAP III.

RNA Polymerases

2. organelle specific RNA polymerases
   more prokaryotic-like
   1. chloroplast
   2. mitochondria
RNA Polymerases

3. RNAP II

   a. core subunits - have sequence similarity to the core subunits of E. coli core RNA polymerase or subunits of other eukaryotic RNA polymerases

   b. shared or common subunits
      same subunits found in RNAP III and II or in RNAP I and RNAP II

   c. unique subunits - no similar homologs found anywhere else
**Figure 20.2** Eukaryotic RNA polymerase II has >10 subunits.

<table>
<thead>
<tr>
<th>Size (kD)</th>
<th>Stoichiometry</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>Related to bacterial subunit binds DNA has CTD = (YSPTSPS)_n [yeast n = 26; mouse n = 52]</td>
</tr>
<tr>
<td>100</td>
<td>1</td>
<td>Related to bacterial subunit binds nucleotides</td>
</tr>
<tr>
<td>50</td>
<td>2</td>
<td>Related to bacterial subunit</td>
</tr>
<tr>
<td>&lt;1</td>
<td>2</td>
<td>Common to all 3 polymerases</td>
</tr>
<tr>
<td>25</td>
<td>1</td>
<td>Common to all 3 polymerases</td>
</tr>
<tr>
<td></td>
<td>&lt;1</td>
<td>Common to all 3 polymerases</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>Common to all 3 polymerases</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Common to all 3 polymerases</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>Common to all 3 polymerases</td>
</tr>
</tbody>
</table>
RNA Polymerases

3. RNAP II
   d. CTD or C-terminal domain of the largest subunit
      i. a heptapeptide repeat - 52X humans and 26X yeast
      ii. can be highly phosphorylated at Ser and Thr
      iii. three subforms of RNAP II: IIo, IIa, and IIb
   e. cycling of phosphorylated and dephosphorylated forms of RNAP II associated with different stages in transcription
   f. CTD is also required to recruit proteins for capping of 5'-end of mRNA, as well as for splicing and polyadenylation of the 3'end of mRNA
Figure 20.13  Phosphorylation of the CTD by the kinase activity of TFIIH may be needed to release RNA polymerase to start transcription.
FIGURE 31.21 X-Ray structure of an RNAP II elongation complex. (a) The RNA · DNA complex in the structure with the template DNA cyan, the non-template DNA green, and the newly synthesized RNA red. The magenta dot marked Mg$^{2+}$ represents the strongly bound active site metal ion. The black box encloses those portions of the complex that are clearly visible in the structure; the double-stranded portion of the DNA marked "Downstream DNA duplex" is poorly ordered, and the remaining portions of the complex are disordered. (b) View of the transcribing complex from the bottom of Fig. 31.20a in which portions of Rpb2 that form the near side of the cleft have been removed to expose the bound RNA · DNA complex. The protein is represented by its backbone in which the clamp, which is closed over the downstream DNA duplex, is yellow, the bridge helix is green, and the remaining portions of the protein are gray. The DNA and RNA are colored as in Part a with their well-ordered portions drawn in ladder form and their less ordered portions drawn in backbone form. The active site Mg$^{2+}$ ion is represented by a magenta sphere. (c) Cutaway schematic diagram of the transcribing complex in Part b in which the cut surfaces of the protein are light gray, its remaining surfaces are darker gray, and several of its functionally important structural features are labeled. The DNA, RNA, and active site Mg$^{2+}$ ion are colored as in Part a with portions of the DNA and RNA that are not visible in the X-ray structure represented by dashed lines. The α-amanitin binding site is marked by an orange dot. [Modified from diagrams by Roger Kornberg, Stanford University. PDBid 1BH]
B. mRNA genes transcribed by RNAP II
   1. TATA box element - located between -30 and -20 bps
   2. Initiator region or Inr: centered on the start site of transcription
   3. DPE: downstream promoter element
   4. Response elements (RE)
      a. upstream of the TATA box
      b. many different kinds - help respond to signals
      c. multiple RE present - synergy
Figure 20.1 Overview: a typical gene transcribed by RNA polymerase II has a promoter that extends upstream from the site where transcription is initiated. The promoter contains several short (<10 bp) sequence elements that bind transcription factors, dispersed over >200 bp. An enhancer containing a more closely packed array of elements that also bind transcription factors may be located several kb distant. (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.)
Figure 20.16 Saturation mutagenesis of the upstream region of the \( \beta \)-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.
Figure 20.17 Promoters contain different combinations of TATA boxes, CAAT boxes, GC boxes, and other elements.
B. mRNA genes transcribed by RNAP II

5. enhancers
   a. can be located at great distances (>1000 bps) from start site of transcription either from the 5' or 3' end of gene
   b. stimulates transcription (~100 times)
   c. orientation independent
   d. two models of how enhancers might work
      i. entry point of RNAP II by preventing nucleosomes from binding or an altered DNA conformation that promotes RNAP II recognition
      ii. transcription factors bound to enhancer will stimulate binding of RNAP II to promoter regions closer to the start site of transcription
Figure 20.20 An enhancer may function by bringing proteins into the vicinity of the promoter. An enhancer does not act on a promoter at the opposite end of a long linear DNA, but becomes effective when the DNA is joined into a circle by a protein bridge. An enhancer and promoter on separate circular DNAs do not interact, but can interact when the two molecules are catenated.
Transcription Factors

General versus promoter specific transcription factors.

Factors that are required for all mRNA genes and others that are required for only a small subset of genes.
Figure 21.1 The regulatory region of a human metallothionein gene contains regulator elements in both its promoter and enhancer. The promoter has elements for metal induction; an enhancer has an element for response to glucocorticoid. Promoter elements are shown above the map, and proteins that bind them are indicated below.

Response elements

- GRE
- BLE
- MRE MRE
- BLE TRE
- MRE GC
- MRE
- TATA

Protein binding

- Steroid-receptor
- AP2
- ?
- AP2
- AP1
- ?
- SP1
- ?

General Transcription Factors

Promoter Specific Transcription Factors