Steps in gene expression: comparison of prokaryotic and eukaryotic cells
Six steps at which eukaryotic gene expression can be controlled.
In prokaryotic cells, genes do not have introns (no step 2) and transcription and translation are not separated in space and time (no step 3).
Transcription of DNA into RNA by RNA polymerase.

1. Requires DNA template, 4 ribonucleotide 5’ triphosphates, Mg$^{+2}$.

2. De novo synthesis: does not require a primer. Low fidelity compared to DNA polymerase: errors $1/10^4$-$10^5$. RNA polymerase incorporates ~30 nt/s (much slower than DNA polymerase).

3. Activity highly regulated in vivo: at initiation, elongation and termination.

4. The nucleotide at the 5’ end of an RNA strand retains all three of its phosphate groups; all subsequent nucleotides release pyrophosphate (PPi) when added to the chain and retain only their α phosphate (red).

5. The released PPi is subsequently hydrolyzed by pyrophosphatase to Pi, driving the equilibrium of the overall reaction toward chain elongation.

6. In most cases, only one DNA strand is transcribed into RNA.
Transcription of two genes on different template strands. Note that growth of the transcript always occurs in the 5’-to-3’ direction.

The convention for designating the 5’ and 3’ ends of a gene
Biochemical studies of bacterial RNA polymerase

1. DNA binding assays- DNaseI footprinting
   Gel shift

2. Role of individual subunits- dissociation of holoenzyme
   by column chromatography
DNase I footprinting, a common technique for identifying protein-binding sites in DNA.

1. A DNA fragment is labeled at one end with $^{32}$P (red dot).

2. Portions of the sample then are digested with DNase I in the presence and absence of a protein that binds to a specific sequence in the fragment.

3. A low concentration of DNase I is used so that on average each DNA molecule is cleaved just once (vertical arrows).

4. The two samples of DNA then are separated from protein, denatured to separate the strands, and electrophoresed. The resulting gel is analyzed by autoradiography, which detects only labeled strands and reveals fragments extending from the labeled end to the site of cleavage by DNase I.
DNaseI footprint of RNAP on the *lac* promoter

Lanes 1 and 2 are DNA sequencing reactions for orientation

Lane 3 is the no protein control

Lane 4 contained RNAP

Footprint:
Advantage- can assay short-lived interactions; proteins have very characteristic binding patterns

Disadvantage- requires nearly stoichiometric binding
“Gel shift”: electrophoretic mobility shift assay ("EMSA") for DNA-binding proteins

1. Prepare labeled DNA probe
2. Bind protein
3. Native gel electrophoresis

Advantage: sensitive
Disadvantage: requires stable complex; little "structural" information about which protein is binding
Ion-exchange chromatography
Dissociation of RNAP and purification of [□] by ion-exchange chromatography

Carboxymethyl- (-CO₂⁻²) or phospho- (-PO₃⁻²) cellulose

[NaCl]

[protein]
**E. coli** RNA polymerase holoenzyme bound to DNA.

<table>
<thead>
<tr>
<th>Subunit</th>
<th>Stoichiometry in holoenzyme</th>
<th>Role</th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>2</td>
<td>Binds regulatory sequences/proteins</td>
</tr>
<tr>
<td>β</td>
<td>1</td>
<td>Polymerase; Binds DNA template</td>
</tr>
<tr>
<td>β’</td>
<td>1</td>
<td>Polymerase; Binds DNA template</td>
</tr>
<tr>
<td>σ70</td>
<td>1</td>
<td>Promoter recognition</td>
</tr>
<tr>
<td>−35</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>−10</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>+1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>DNA</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Direction of transcription**

**Upstream, downstream**
Thermus aquaticus core RNA polymerase

PDB (Protein Data Bank)
http://www.rcsb.org/pdb/

File 1HQM
A single RNA polymerase makes multiple types of RNAs in prokaryotes

<table>
<thead>
<tr>
<th>Type</th>
<th>Relative amount (%)</th>
<th>Sedimentation coefficient (S)</th>
<th>Mass (kd)</th>
<th>Number of nucleotides</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ribosomal RNA (rRNA)</td>
<td>80</td>
<td>23</td>
<td>$1.2 \times 10^3$</td>
<td>3700</td>
</tr>
<tr>
<td></td>
<td></td>
<td>16</td>
<td>$0.55 \times 10^3$</td>
<td>1700</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>$3.6 \times 10^1$</td>
<td>120</td>
</tr>
<tr>
<td>Transfer RNA (tRNA)</td>
<td>15</td>
<td>4</td>
<td>$2.5 \times 10^1$</td>
<td>75</td>
</tr>
<tr>
<td>Messenger RNA (mRNA)</td>
<td>5</td>
<td></td>
<td>Heterogeneous</td>
<td></td>
</tr>
</tbody>
</table>
The dissociable sigma subunit gives promoter specificity to prokaryotic RNA polymerase (RNAP)

Core enzyme

No specific promoter binding; tight non-specific DNA binding ($K_d \sim 5 \times 10^{-12} \text{ M}$)

Holoenzyme

Specific promoter binding; weak non-specific DNA binding ($K_d \sim 10^{-7} \text{ M}$)
Transcription initiation by prokaryotic RNA polymerase

Holoenzyme  "scanning"  Promoter

Closed complex

rNTPs  PPI

Open complex; initiation

5’pppA  mRNA
### Sigma Factors of *E. coli*

<table>
<thead>
<tr>
<th>Sigma Factor</th>
<th>Promoters Recognized</th>
<th>Promoter Consensus</th>
</tr>
</thead>
<tbody>
<tr>
<td>σ^70</td>
<td>Most genes</td>
<td>(-35) Region: TTGACAT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-10) Region: TATAAT</td>
</tr>
<tr>
<td>σ^32</td>
<td>Genes induced by heat shock</td>
<td>(-35) Region: TCTCNCCCTTGAA CCCCATNTA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-10) Region: CCGATAT</td>
</tr>
<tr>
<td>σ^28</td>
<td>Genes for motility and chemotaxis</td>
<td>(-35) Region: CTAAA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-10) Region: CCGATAT</td>
</tr>
<tr>
<td>σ^38</td>
<td>Genes for stationary phase and stress response</td>
<td>(-24) Region: ?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-12) Region: ?</td>
</tr>
<tr>
<td>σ^54</td>
<td>Genes for nitrogen metabolism and other functions</td>
<td>(-24) Region: CTGGNA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-12) Region: TTGCA</td>
</tr>
</tbody>
</table>
Transcriptional elongation
Supercoiling of DNA during transcription causes a requirement for topoisomerase.
Rho-independent prokaryotic transcription termination

Figure 25-10. A hypothetical strong (efficient) E. coli terminator. The base sequence was deduced from the sequences of several transcripts. (a) The DNA sequence together with its corresponding RNA. The A·T-rich and G·C-rich sequences are shown in blue and red, respectively. The twofold symmetry axis (green symbol) relates the flanking shaded segments that form an inverted repeat. (b) The RNA hairpin structure and poly(U) tail that trigger transcription termination. [After Pribnow, D., in Goldberger, R.F. (Ed.), Biological Regulation and Development, Vol. 1, p. 253, Plenum Press (1979).]
Rho-dependent transcription termination

Rho: forms RNA-dependent hexameric ATPase, translocates along RNA 5’-to-3’, promoting dissociation from polymerase