Transcription activation in multicellular organisms.

Liam.Keegan@ceitec.muni.cz

Eukaryotic total RNA.

Ribosomal RNAs and tRNAs are major bands, mRNA is a smear on denaturing gel stained with Ethidium Bromide.

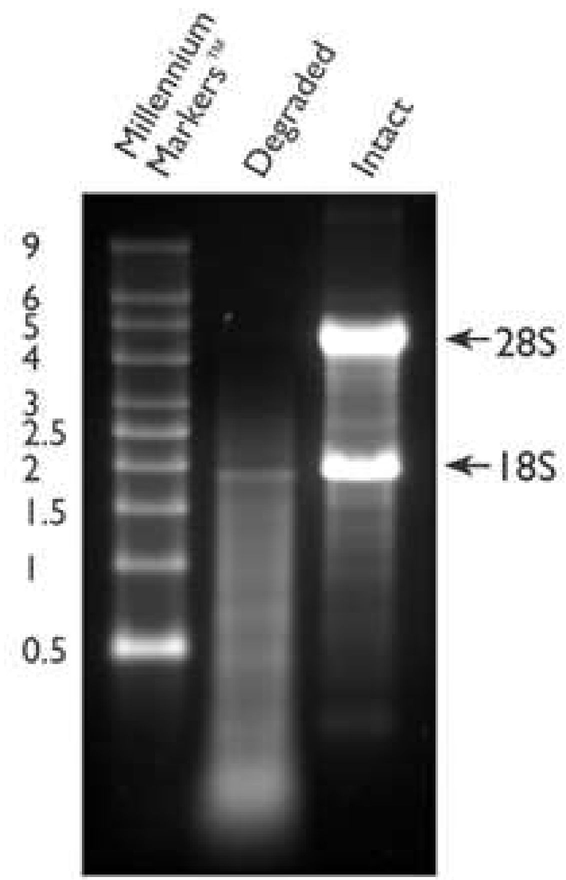
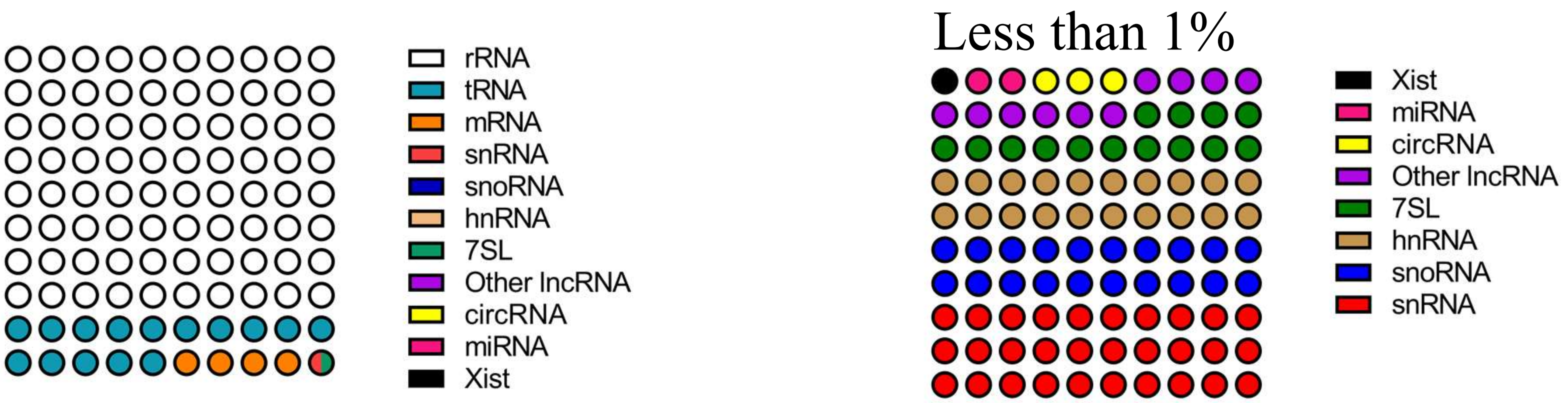


Figure 1. Intact vs. Degraded RNA. Two µg of degraded total RNA and intact total RNA were run beside Ambion's RNA Millennium Markers™ on a 1.5% denaturing agarose gel. The 18S and 28S ribosomal RNA bands are clearly visible in the intact RNA sample. The degraded RNA appears as a lower molecular weight smear.

Total RNA

RNA by mass



RNA by number of molecules

Adapted from doi: 10.3389/fgene.2015.00002

• Eukaryotic RNA Polymerases

- Three nuclear RNA polymerases (*E. coli* has only one).
- RNA pol I (in nucleolus)
- transcribes rRNA genes → 50 – 70 % cell's RNA synthesis
- resistant to > 500 μg/ml α-amanitin, an octapeptide from
- *Amanita phalloides* (Death Cap Mushroom) that grows near Oak trees.
- RNA pol II (in nucleoplasm)
- transcribes all protein-encoding genes & most small nuclear RNAs
- → 20 – 40% cell's RNA synthesis
- inhibited by low ~ 0.03 μg/ml α-amanitin
- RNA pol III (in nucleoplasm)
- transcribes 5S, tRNA genes & some small nuclear RNAs
- → < 10% cell's RNA synthesis
- inhibited by 20 μg/ml α-amanitin in animal cells
- resistant to α-amanitin in yeast and insects

20.2 Eukaryotic RNA Polymerases Consist of Many Subunits

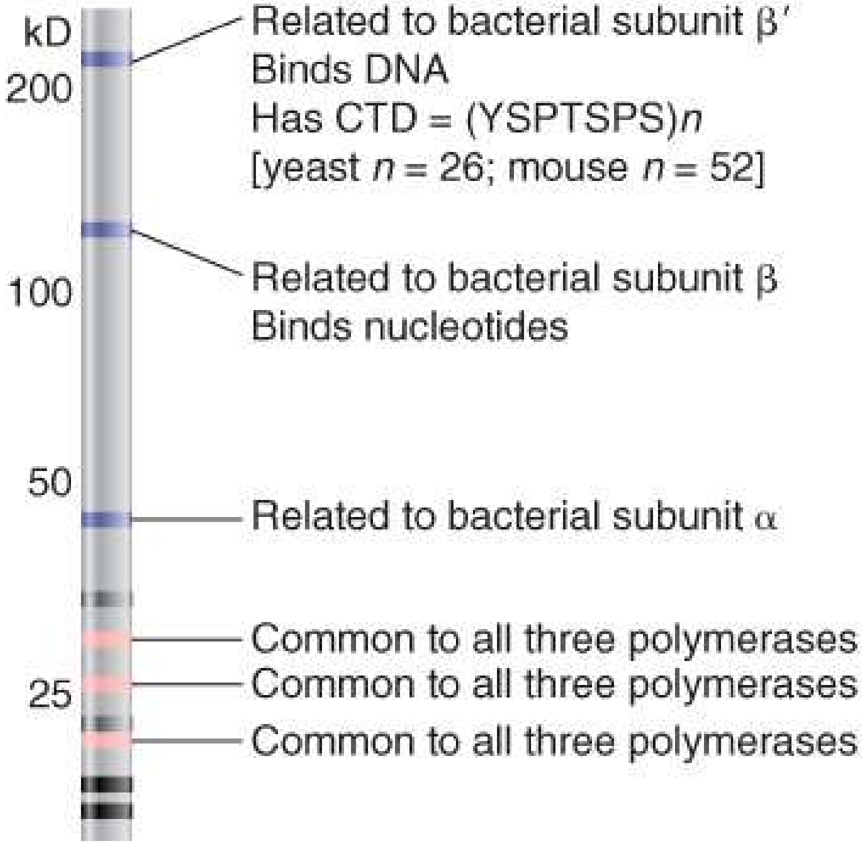
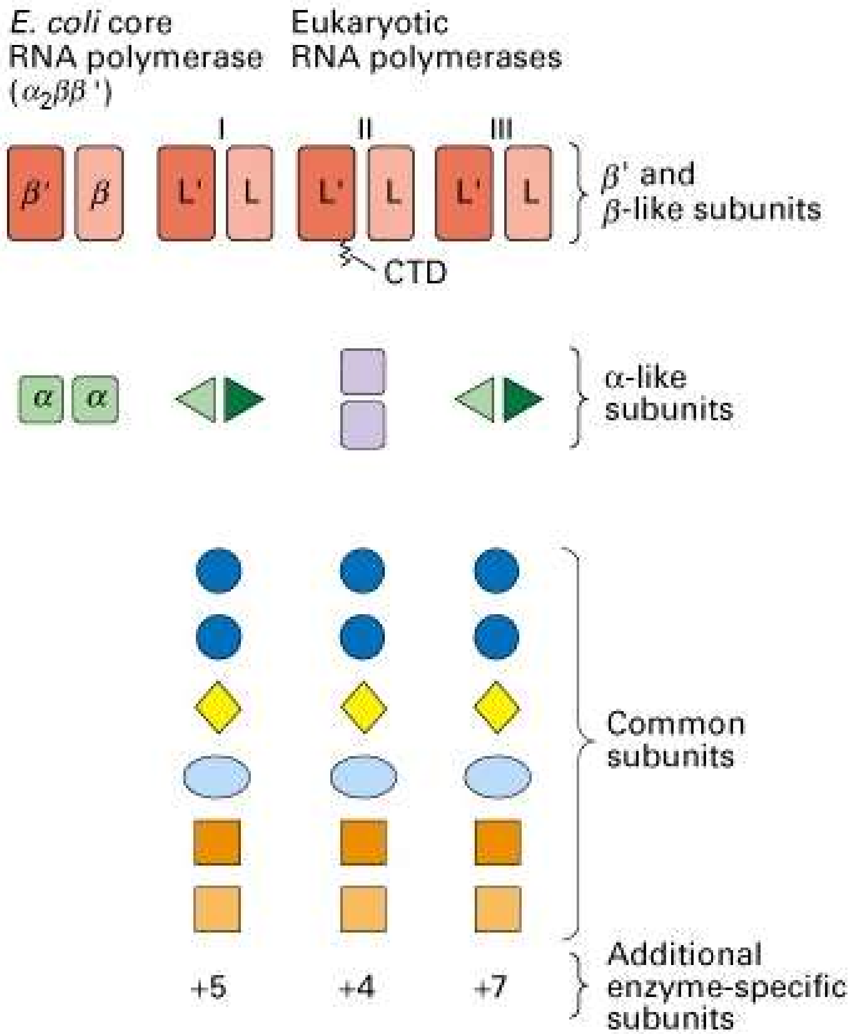


Figure 20.02: Some subunits are common to all classes of eukaryotic RNA polymerases and some are related to bacterial RNA polymerase.

- All eukaryotic RNA polymerases have ~12 subunits and are complexes of ~500 kD.
- Some subunits are common to all three RNA polymerases.
- The largest subunit in RNA polymerase II has a **CTD (carboxy-terminal domain)** consisting of multiple repeats of a heptamer.

Eukaryotic RNA polymerases are similar to that of *E. coli* but have 12 subunits.



Transcription of protein-coding mRNAs by RNA polymerase II.

Defining RNA polymerase II promoters.

20.5 The Start Point for RNA Polymerase II

- RNA polymerase II requires general transcription factors (called $TF_{II}X$) to initiate transcription.
- RNA polymerase II promoters frequently have a short conserved sequence Py_2CAPy_5 (the **initiator Inr**) at the start point.
- The **TATA box** is a common component of RNA polymerase II promoters and consists of an A-T-rich octamer located ~25 bp upstream of the start point.

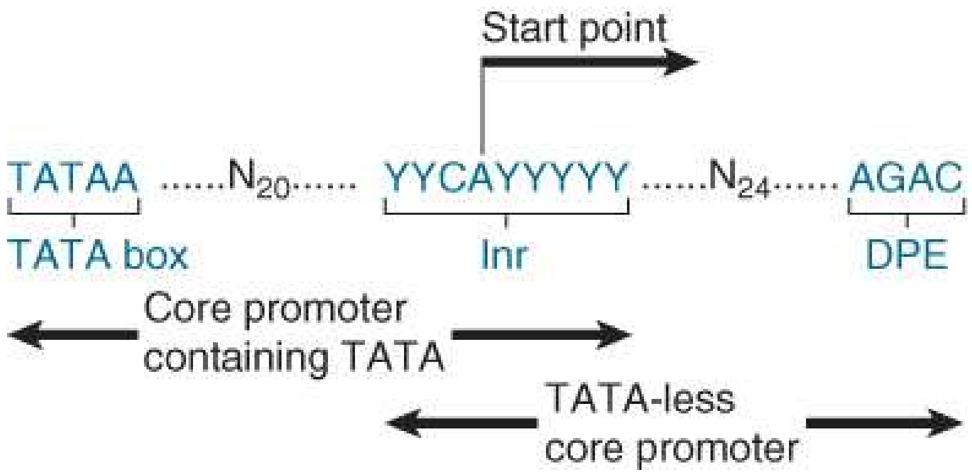
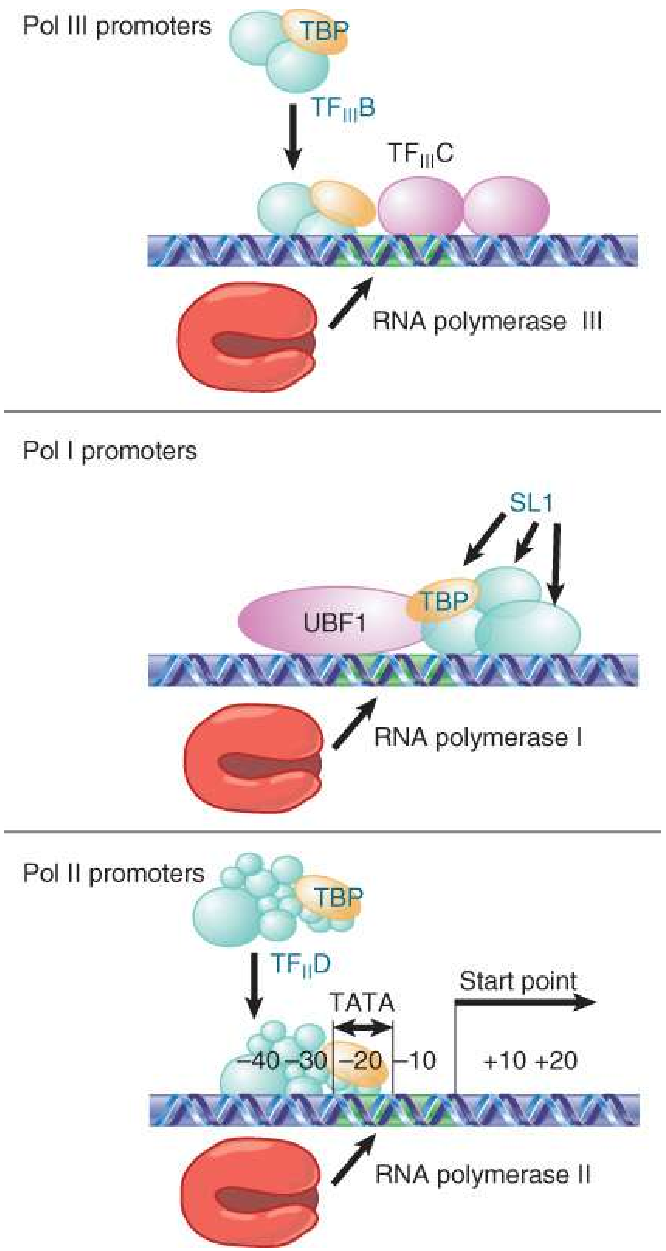


Figure 20.07: A minimal pol II promoter may have a TATA box ~25 bp upstream of the Inr.

20.5 The Start Point for RNA Polymerase II

- The **downstream promoter element (DPE)** is a common component of RNA polymerase II promoters that do not contain a TATA box (**TATA-less promoters**).
- A core promoter for RNA polymerase II includes the Inr and, commonly, either a TATA box or a DPE.
  - It may also contain other minor elements.

20.6 General transcription factors (GTFs) TBP in TFIID Is a Universal Factor



- TATA-binding protein (TBP) is a component of each of the different positioning factors 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 ~14 **TAFs**, (TBP-associated factors) with a total mass ~800 kD.

Figure 20.08: RNA polymerases are positioned at all promoters by a factor that contains TBP.

20.6 TBP Is a Universal Factor

- TBP binds to the TATA box in the minor groove of DNA.
- TBP forms a saddle around the DNA and bends it by ~80°.

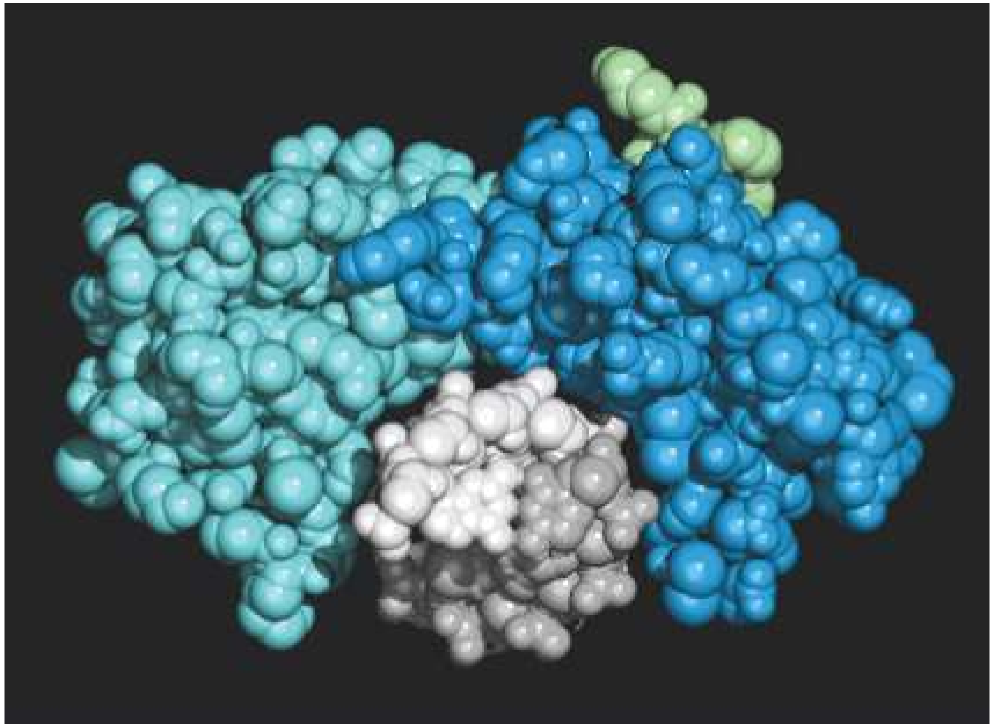


Figure 20.09: A view in cross-section shows that TBP surrounds DNA from the side of the narrow groove.

Photo courtesy of Stephen K. Burley

20.7 The Basal Apparatus Assembles at the Promoter

- The upstream elements and the factors that bind to them increase the frequency of initiation.
- Binding of $TF_{II}D$ to the TATA box or Inr is the first step in initiation.

20.7 The Basal Apparatus Assembles at the Promoter

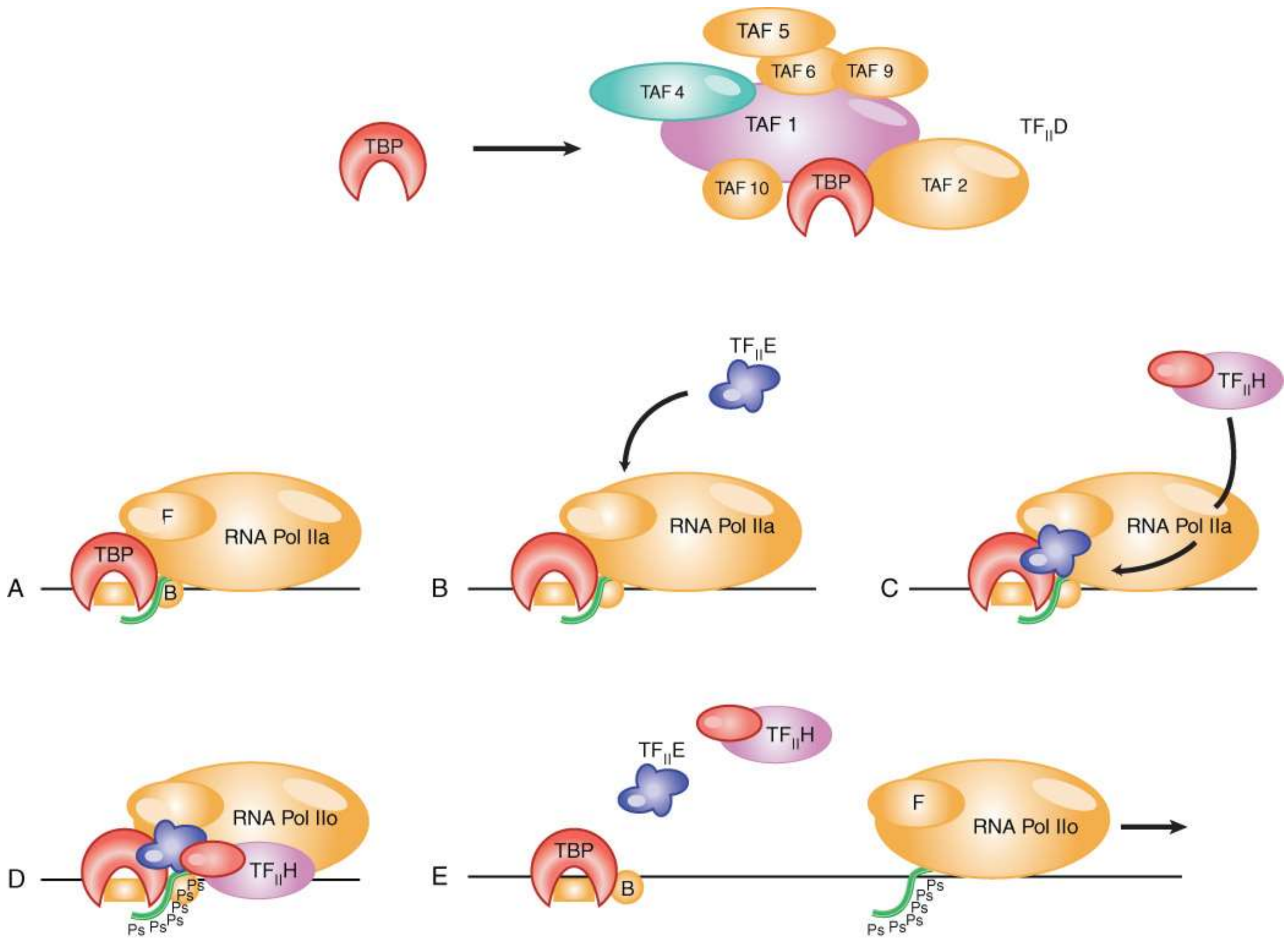


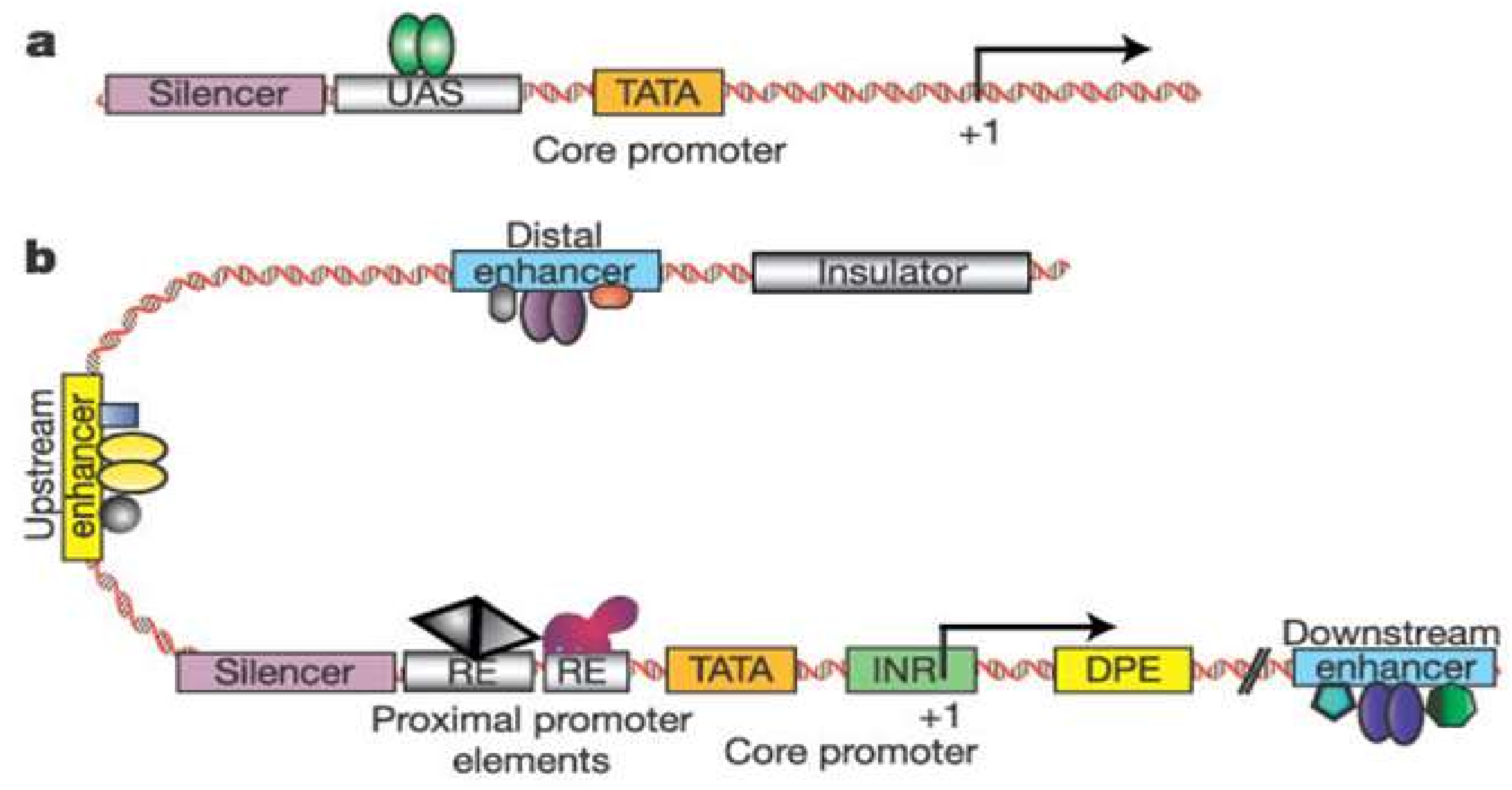
Figure 20.11: An initiation complex assembles at promoters for RNA polymerase II by an ordered sequence of association with transcription factors.

Adapted from M. E. Maxon, J. A. Goodrich, and R. Tijan,
Genes Dev. 8 (1994): 515-524.

20.8 Initiation Is Followed by Promoter Clearance and Elongation

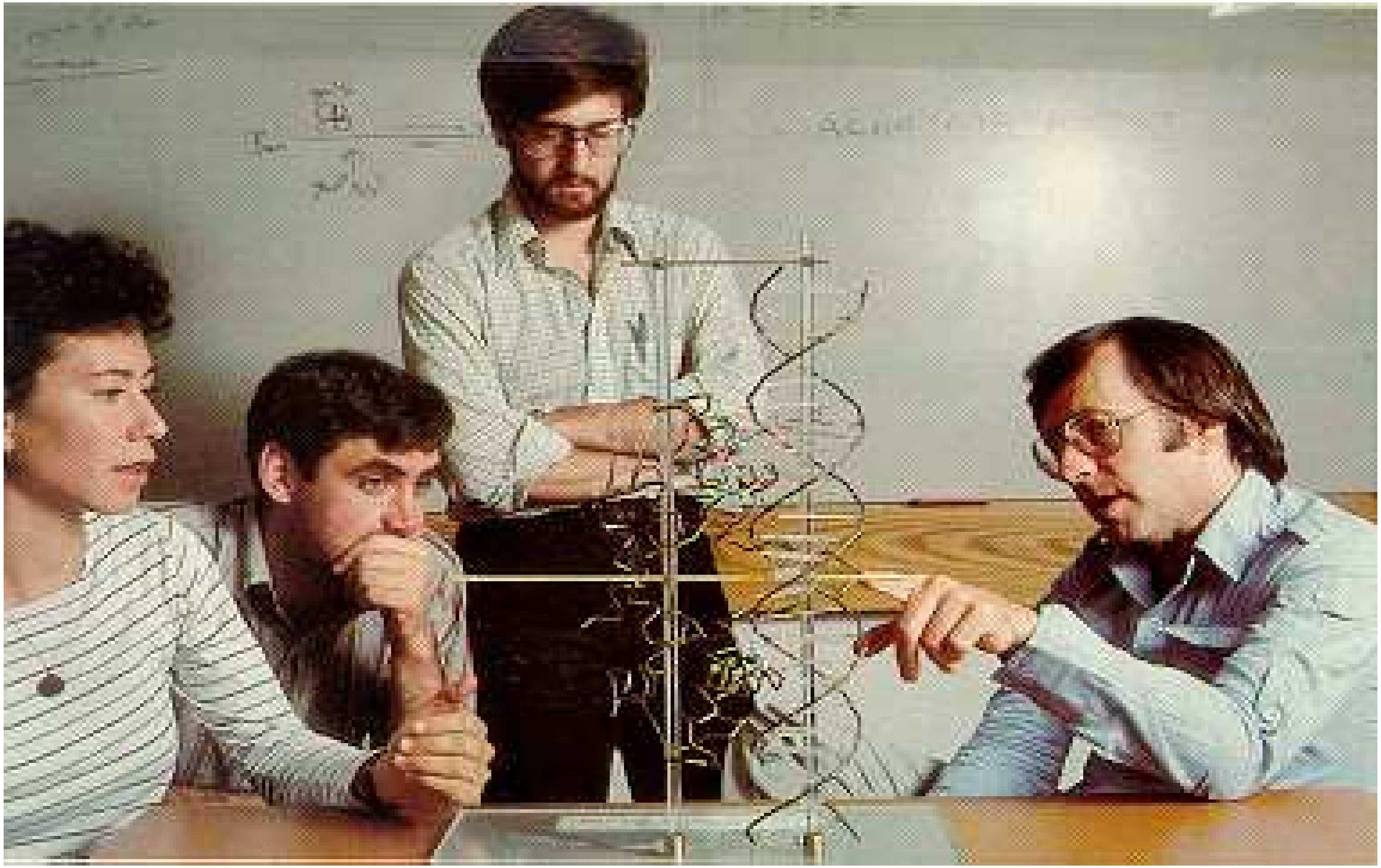
- $TF_{II}B$, $TF_{II}E$, and $TF_{II}H$ are required to melt DNA to allow polymerase movement.
- Phosphorylation of the CTD is required for promoter clearance and elongation to begin.
- Further phosphorylation of the CTD is required at some promoters to end pausing and abortive initiation.

Regulatory elements controlling RNA Pol. II transcription in yeast and higher eukaryotes.



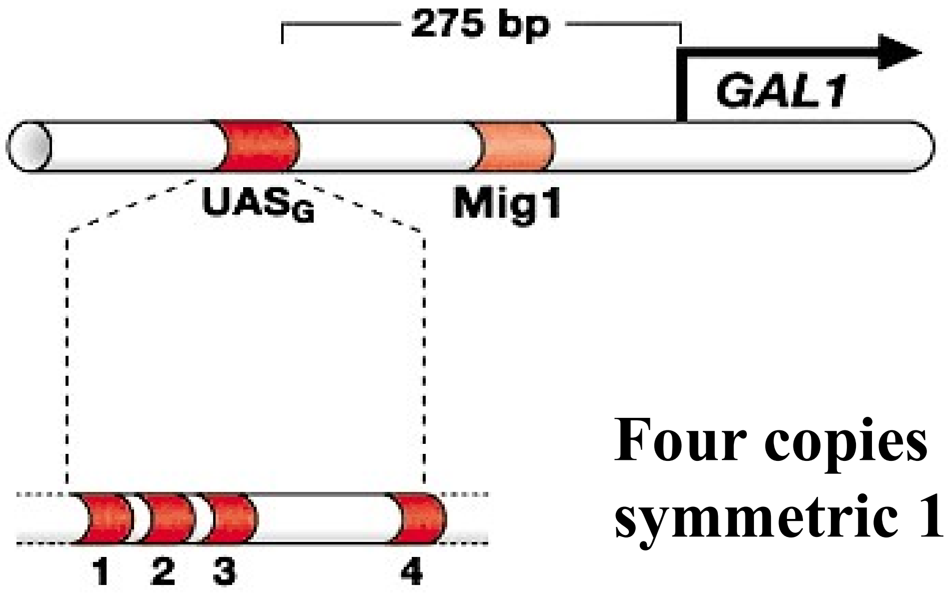
- Promoters
- Proximal regulatory elements
- Enhancers

Specific DNA sequence recognition by DNA-binding activator or repressor proteins is the key to differential gene expression.



Mark Ptashne with graduate students Cynthia Wohlberger, Liam Keegan, Ed Giniger at Harvard, 1982

Deletion analysis of yeast *GAL1* upstream region defined an Upstream Activating Site (UAS_G).



Four copies of a nearly symmetric 17 bp GAL4 site.

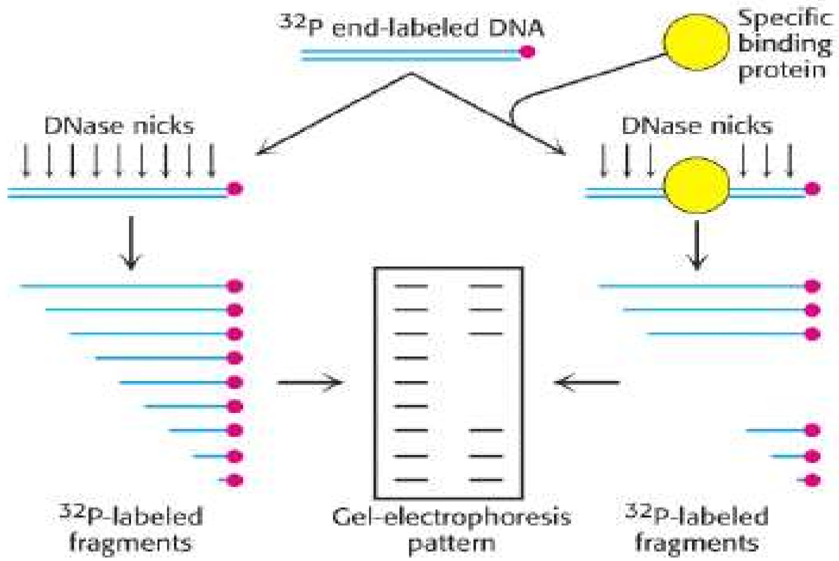
GAL1-lacZ and GAL10-lacZ fusions used for deletion analysis.

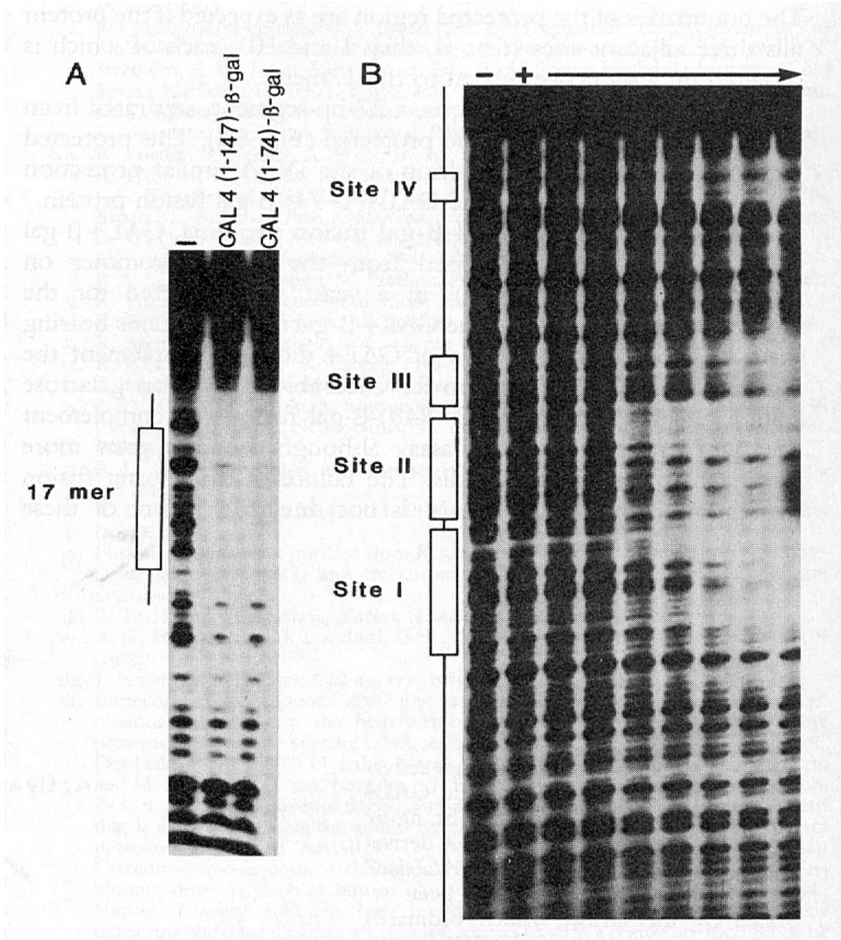
UAS_G had many properties of the Enhancer defined in SV40.

Showing that GAL4 binds the *GAL UAS*.

Methods for identifying and characterising sequence-specific DNA-binding proteins.

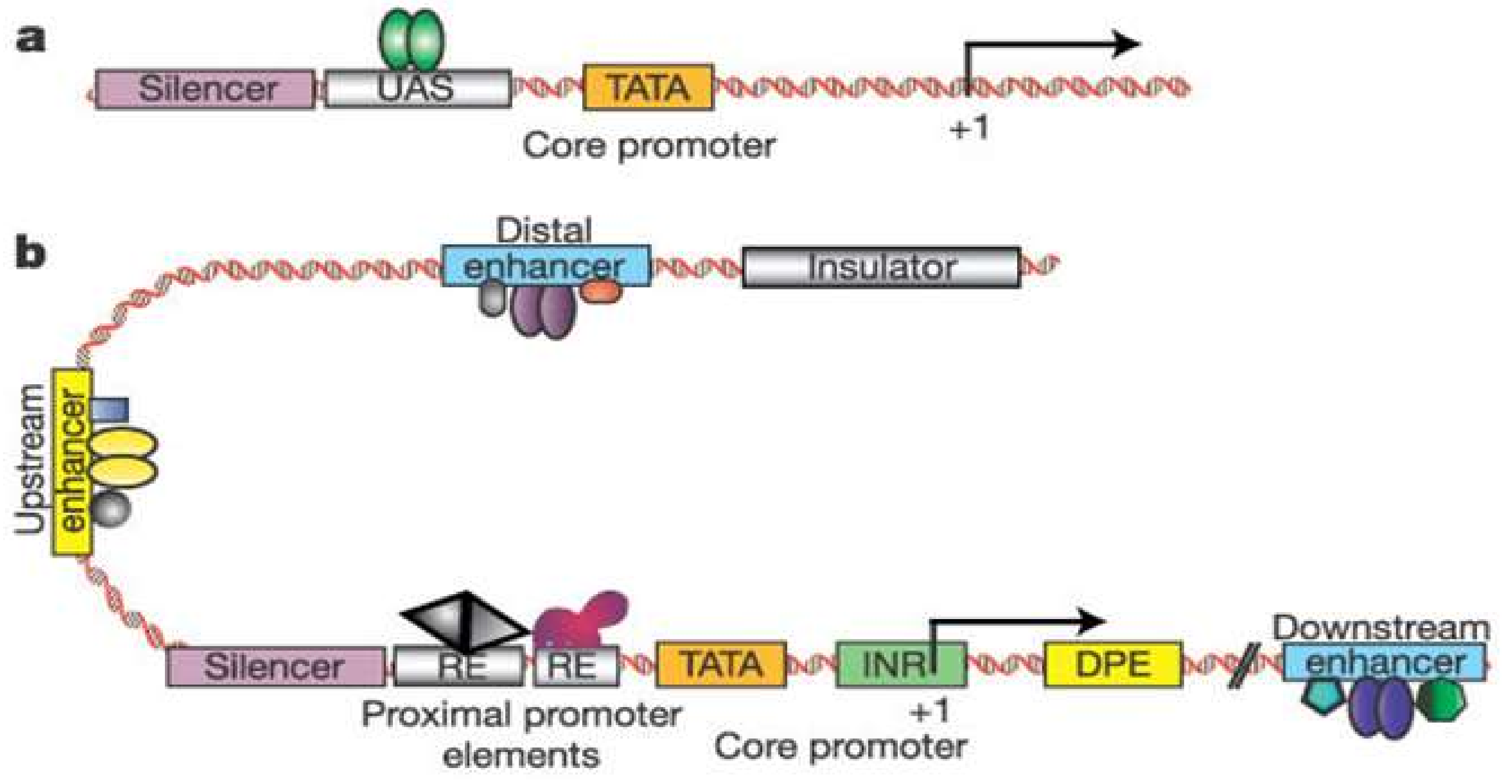
DNase I protection assay to precisely define binding sites of sequence-specific DNA binding proteins and to measure DNA-binding affinity, (DNase I Footprinting)





The GAL4-binding site is CGG-N(11)-CCG.

Regulatory elements controlling RNA Pol. II transcription in yeast and higher eukaryotes.



- Promoters
- Proximal regulatory elements
- Enhancers

25.7 Response Elements Are Recognized by Activators

- Response elements may be located in promoters or enhancers. Proximal regulatory elements were targets of the earliest studies.

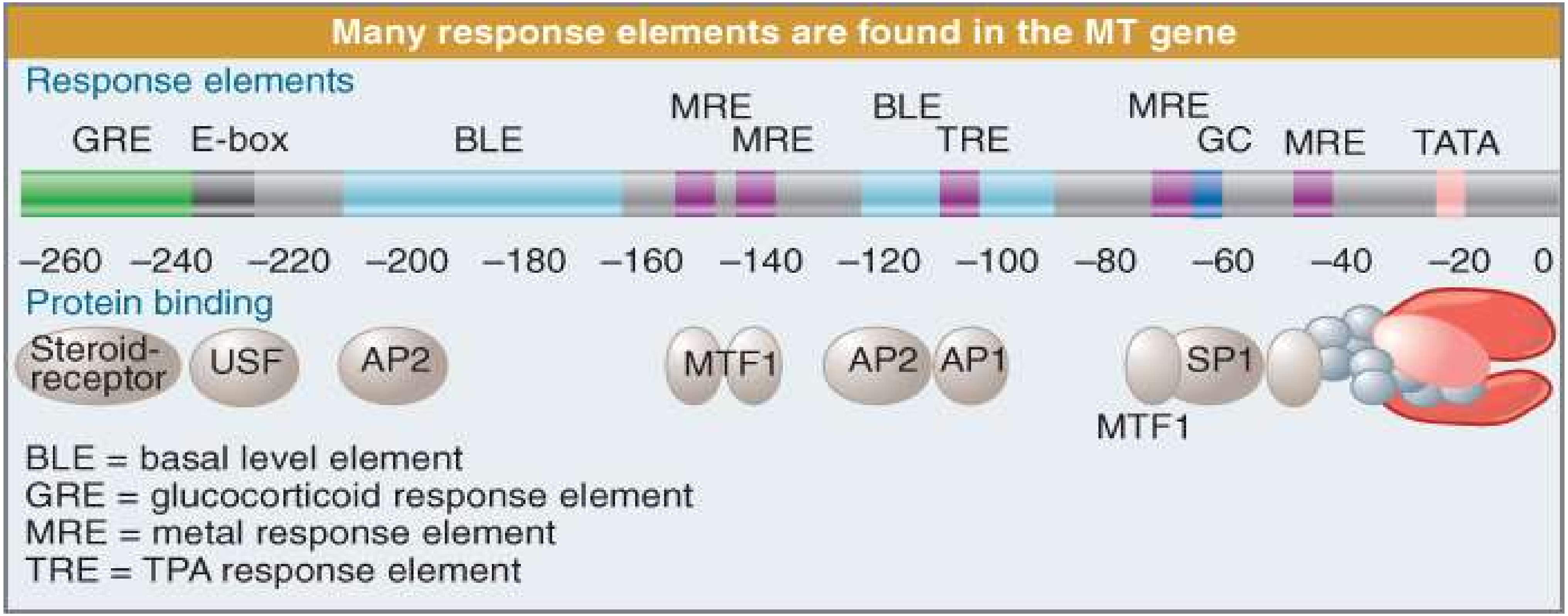
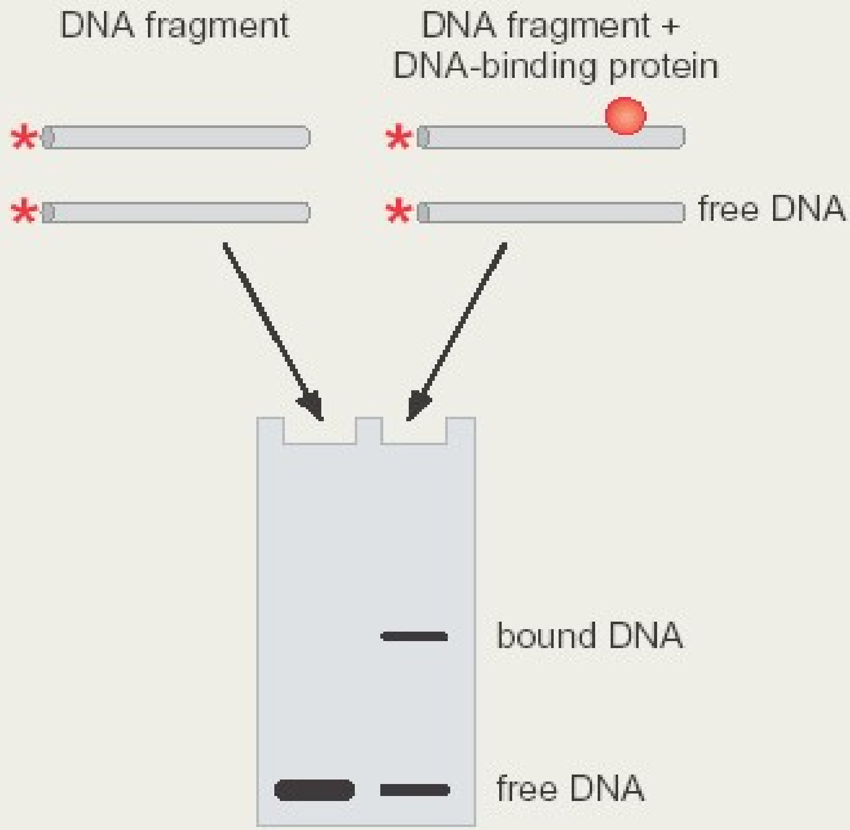


Figure 25.11

Many gene regulators from mammalian cells were identified and purified using the
<u>Bandshift assay or Gel mobility shift assay.</u>

(also Gel Retardation assay, Electrophoretic Mobility Shift Assay (EMSA)).



BOX 16-1 FIGURE 2 Gel mobility shift assay. The principle of the mobility shift assay is shown schematically. A protein is mixed with radiolabeled probe DNA containing a binding site for that protein. The mixture is resolved by acrylamide gel electrophoresis and visualized using autoradiography. DNA not mixed with protein runs as a single band corresponding to the size of the DNA fragment (left lane). In the mixture with the protein, a proportion of the DNA molecules (but not all of them at the concentrations used) binds the DNA molecule. Thus, in the right-hand lane, there is a band corresponding to free DNA, and another corresponding to the DNA fragment in complex with the protein.

Gel is non-denaturing and buffer is mild to avoid disrupting complex.
Gel matrix ‘traps’ protein DNA complex keeping components together to prevent disassociation.

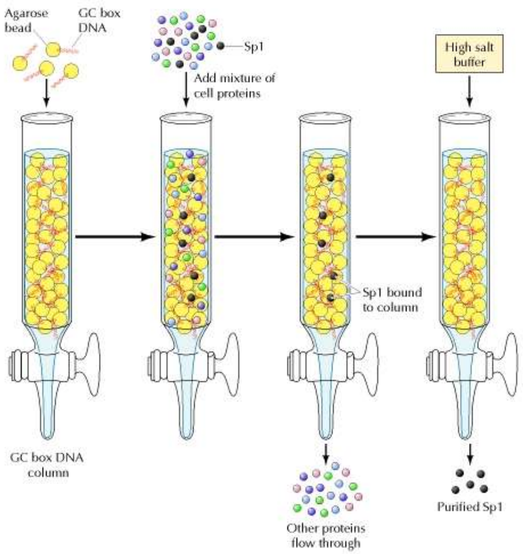
Advantages of the bandshift assay.

- DNase1 footprinting required pure proteins but bandshift assays could detect different scarce DNA-binding proteins even in crude nuclear extracts.

- If the probe DNA fragment comes from an enhancer then a point mutation that inactivates the enhancer should also prevent DNA binding by the key regulatory protein in the extract. (Some proteins bind DNA ends or other sites in the probe DNA)

- The specific bandshift provides a convenient assay to purify the regulatory protein.

DNA-affinity chromatography using a specific oligonucleotide target sequence can purify the binding protein in one step.



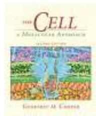
Oligonucleotide attached to affinity column contains repeats of a DNA binding site for Sp1 prote

(5’-GGGCGGG-3’)N

Figure 6.25 Purification of Sp1 by DNA-affinity chromatography

A double-stranded oligonucleotide containing repeated GC box sequences is bound to agarose beads, which are poured into a column. A mixture of cell proteins containing Sp1 is then applied to the column; because Sp1 specifically binds to the GC box oligonucleotide, it is retained on the column while other proteins flow through. Washing the column with high salt buffer then dissociates Sp1 from the GC box DNA, yielding purified Sp1.

From: Regulation of Transcription in Eukaryotes

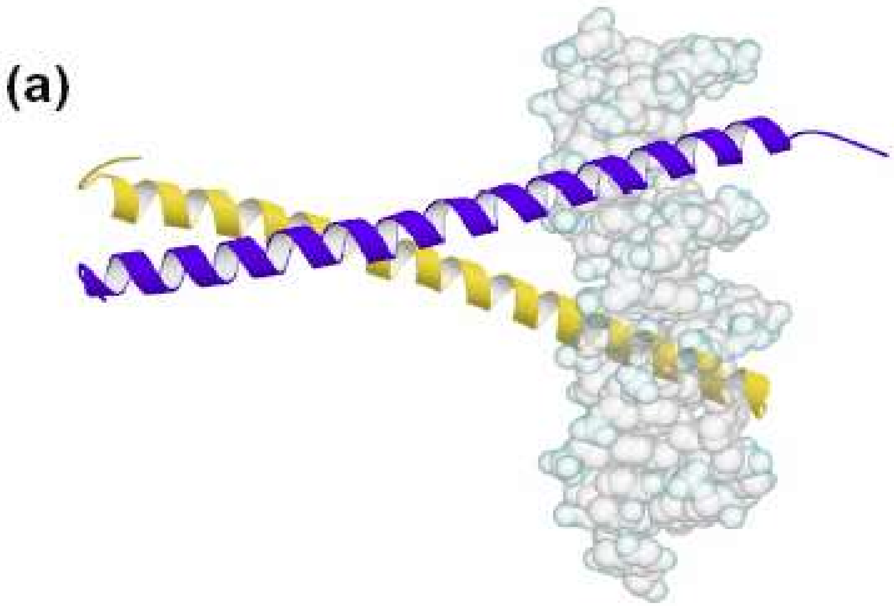


The Cell: A Molecular Approach. 2nd edition.
Cooper GM.
Sunderland (MA): Sinauer Associates; 2000.

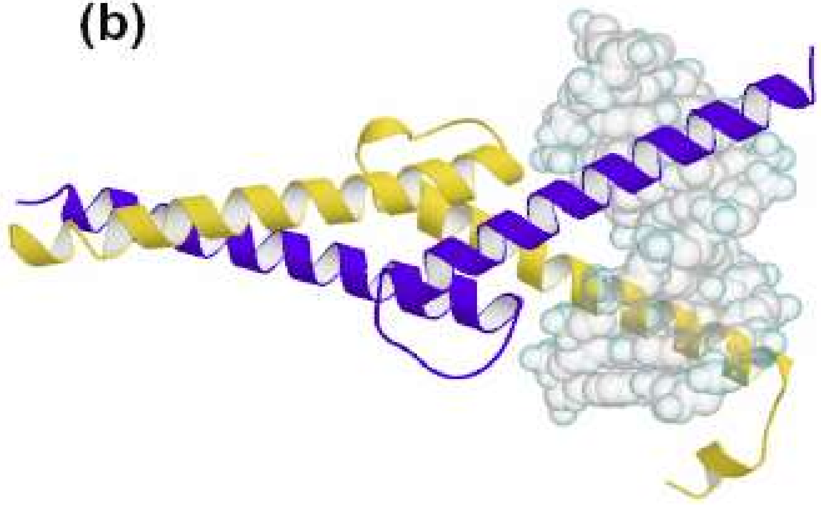
Examples of important transcription regulators identified after purification from nuclear extracts.

- **Fos/jun, AP1**. oncoproteins studied by Tom Curran and others. Steve MacKnight showed these proteins contain a leucine zipper.
- **Nf-kappaB** (Baltimore lab, MIT) Nuclear factor binding the immunoglobulin kappa B enhancer. Important protein in innate immunity and response to viral infection.
- **SRF** (Maniatis lab and Richard Treisman in London) Serum response factor activates genes involved in growth of cells in tissue culture.
- **SP1** (Tjian lab)

The simplest sequence-specific DNA recognition mechanism of all!
CAAT/Enhancer binding protein (C/EBP).



21. Leucine zipper (2dgc)



22. Helix-loop-helix (1am9)

The mechanism of transcription activation by the gene-specific regulators.

- The primary idea is the same as for GAL4 in yeast.
- Recruitment of transcription inititation factors and RNA polymerase to the promoter by DNA loop formation between the enhancer and the promoter.

Mediator.

- *In vitro* transcription extracts were developed in which Pol II and the General Transcription Factors responded with increased transcription when GAL4 or another activator was added.

- This allowed further purifications to isolate additional components needed to mediate the activation. Depending on what activator was used this purified coactivators or mediator complex.

- Some argue that the complex of Pol II, GTFs and Mediator can exist as a ‘holoenzyme’ that recruits to promoters as in *E. coli*.

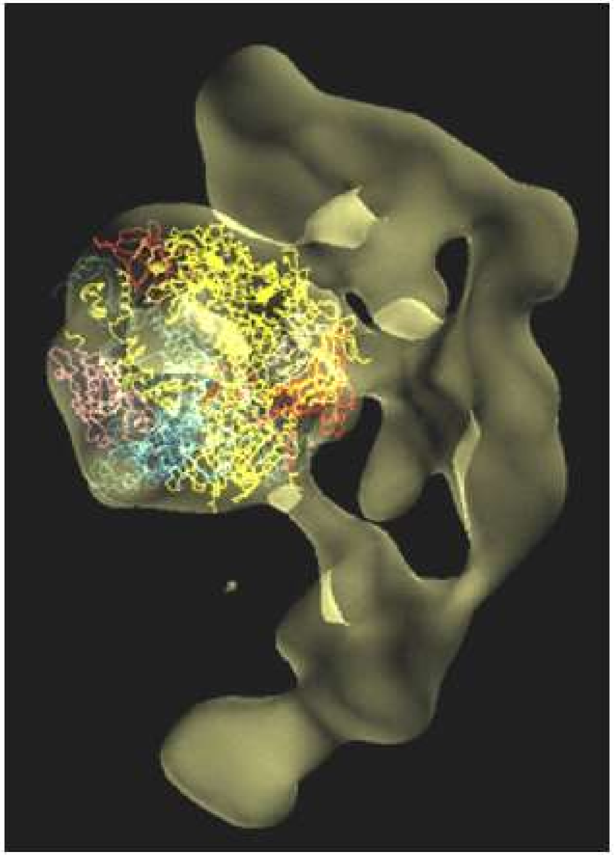
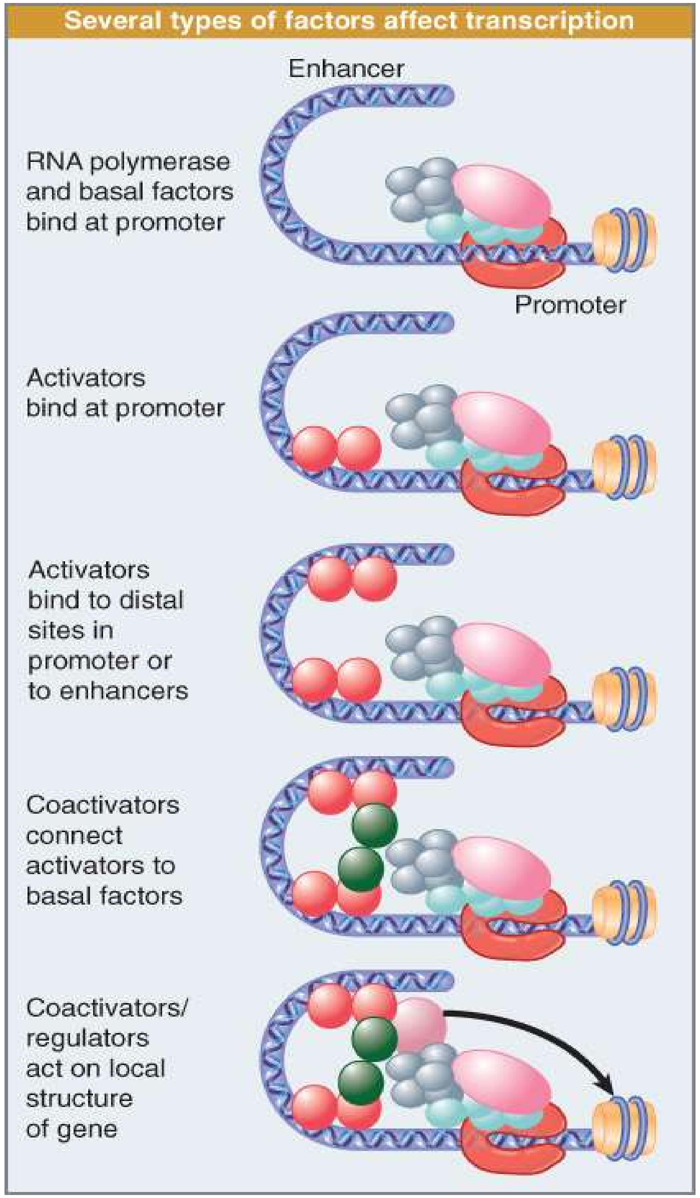


Fig. 18.

Cryo-EM structure of an RNA polymerase II-Mediator complex. The pol II structure was docked in the central density, and is shown in a similar direction of view and color scheme as Fig. 5.

25.2 Gene specific transcription factors (Activators/Coactivators) are distinct from the General Transcription factors .



- **The basal apparatus determines the startpoint for transcription.**
- **Activators determine the frequency of transcription.**
- **Activators work by making protein–protein contacts with the basal factors.**
- **Activators may work via coactivators.**
- **Some components of the transcriptional apparatus work by changing chromatin structure. Coactivator p300/CBP (CREB-binding protein) is a histone acetylase targeting H4.**

Figure 25.2

- Several factors in the basal apparatus are targets with which activators or coactivators interact. Many activators contact different TAFs. Acidic activators contact TFIIB.

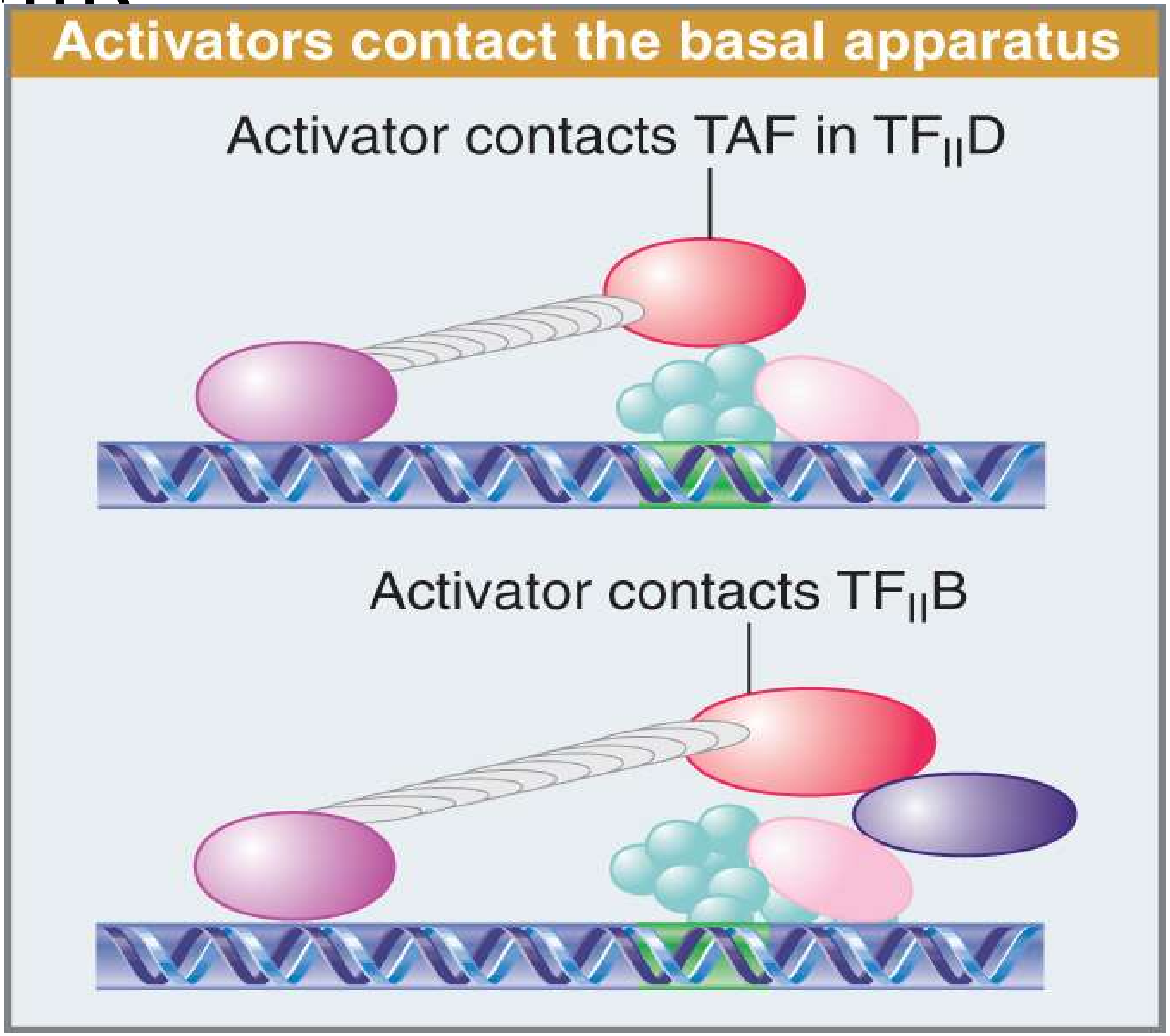


Figure 25.8

- An activator that does not have an activating domain may work by binding a coactivator that has an activating domain. Strong viral activators like Herpesvirus VP16 and Adenovirus E1A are recruited by binding other proteins.

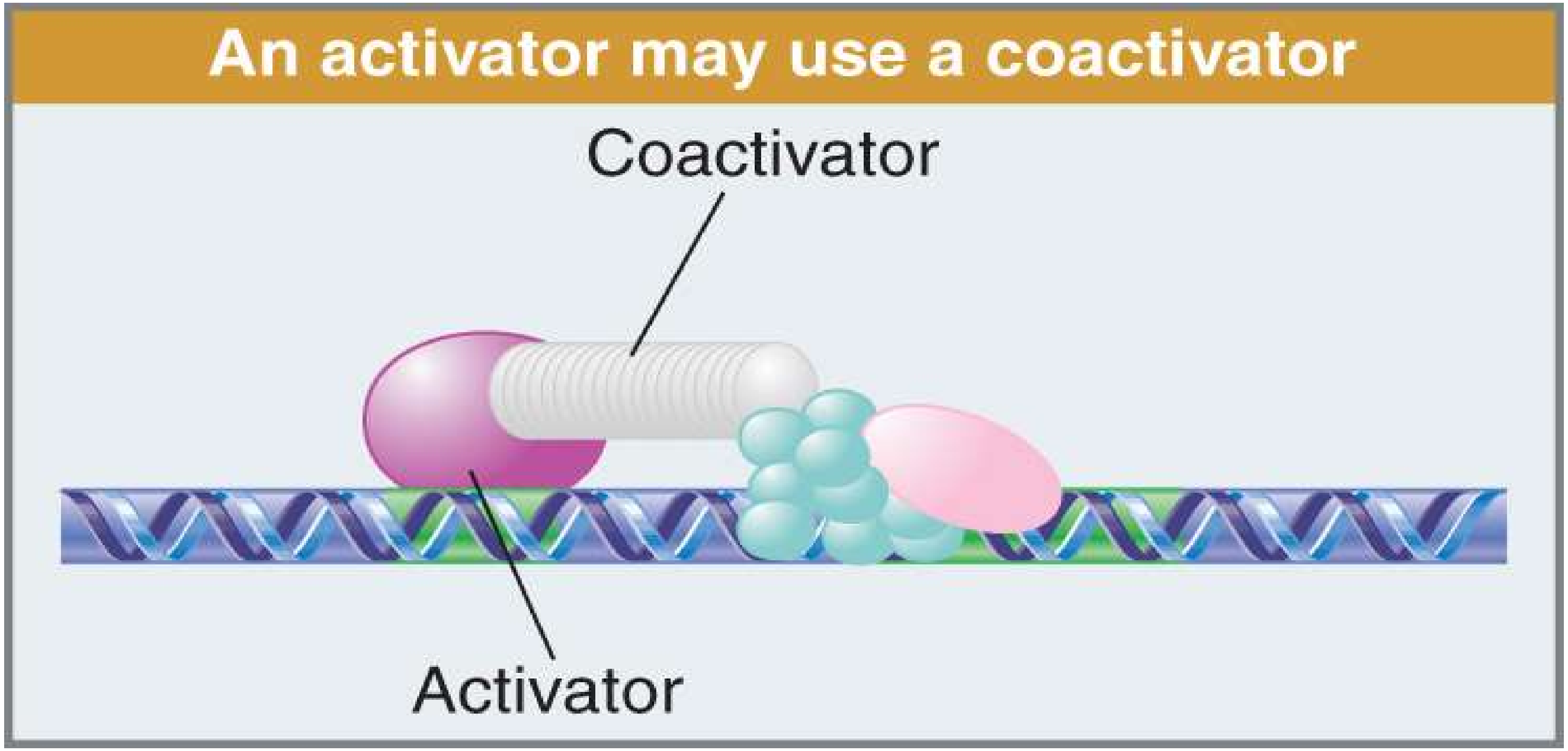
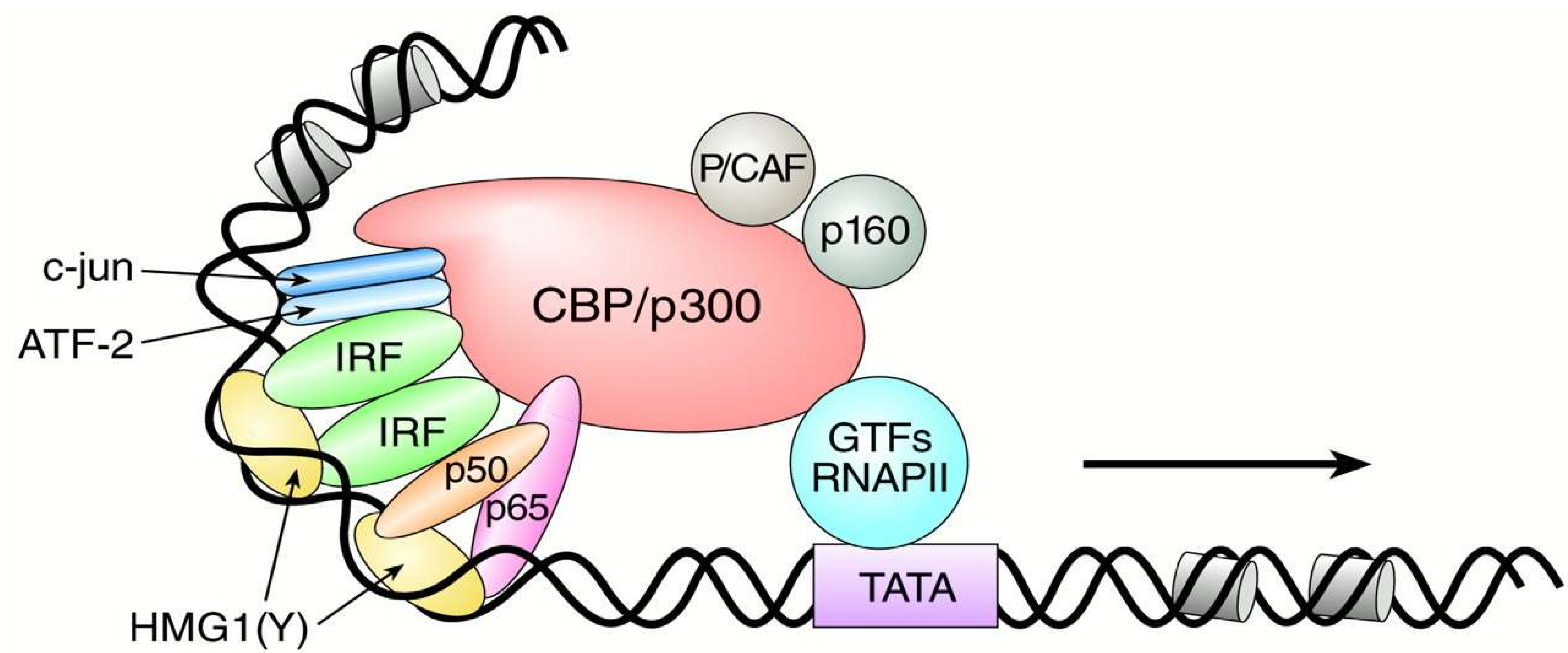


Figure 25.7

The IFN-β enhanceosome complex.



Vo N , Goodman R H J. Biol. Chem. 2001;276:13505-13508

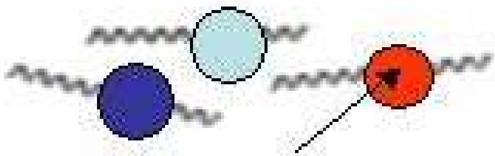
Transcription activation in chromatin.

- Getting transcription initiation and later regulated transcription on chromatin templates required extra factors that operate on nucleosomes.
- Constructs with enhancers reintroduced in mammalian chromosomes were often silenced. This turned out later to be due to lack of regulatory elements like the globin gene Locus Control Region (LCR).
- Only certain sequence-specific DNA binding proteins can still bind their sites if histones are added first (pioneer transcription factors).

Chromatin immunoprecipitation for genome-wide location analysis of chromosomal proteins in living cells

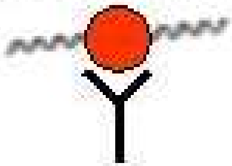
1. Formaldehyde cross link DNA to transcription factors.

2. Lyse cells and fragment chromatin by sonication

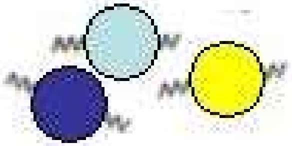


Factor X bound to target gene

3. Immunoprecipitation



Fragments bound to anti-X antibodies
Purify on beads



Unbound fragments are washed away

4. Reverse cross-links, extract DNA

5. Analyse which DNA fragments are enriched in the immunoprecipitated fraction.

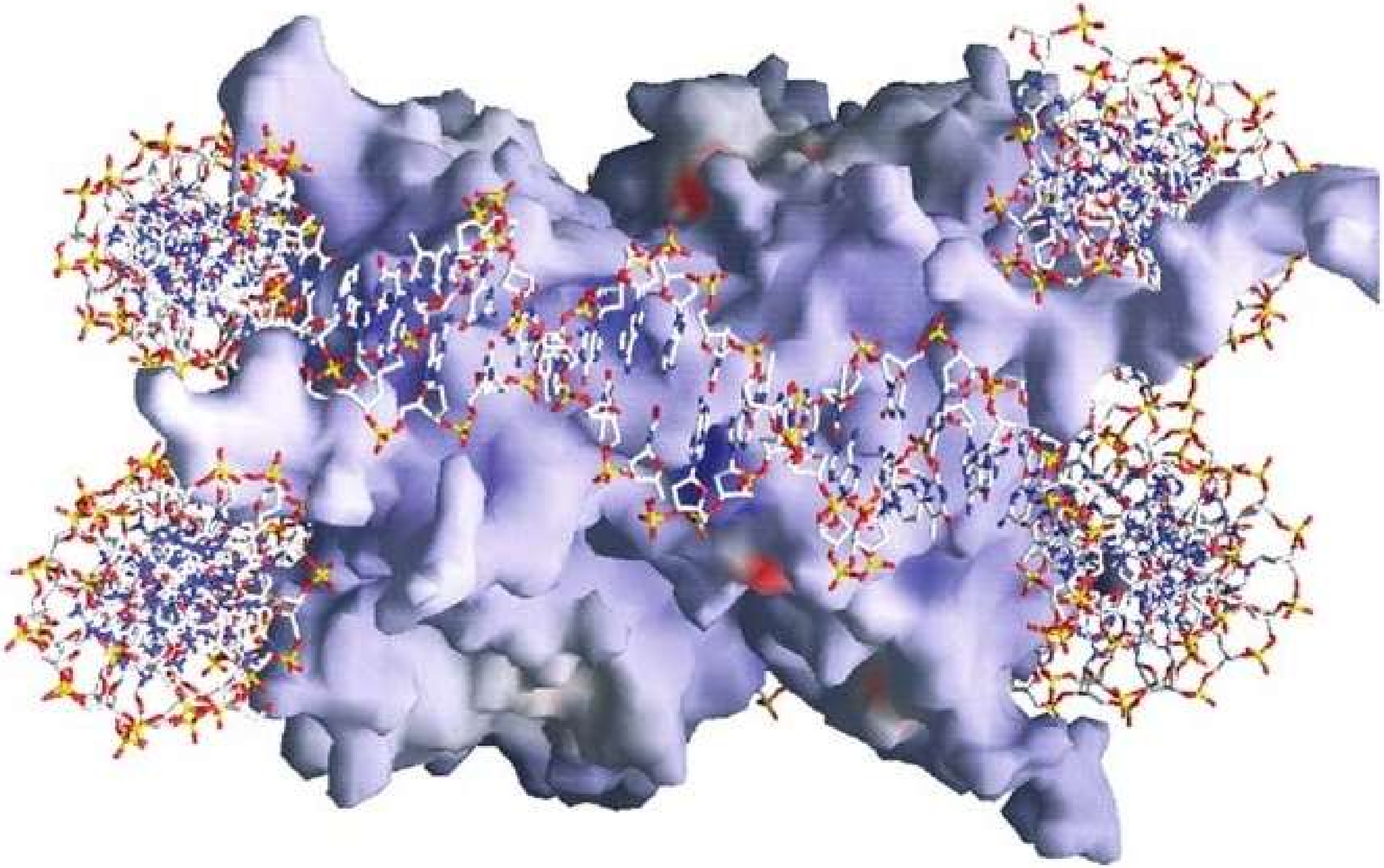
6. Best done now by Next Generation sequencing (ChIP-Seq)

7. Compare reads of immunoprecipitated sequences back to the genome sequence to locate w the protein was bound.

Mapping Polycomb-repressed domains in the bithorax complex using in vivo formaldehyde cross-linked **chromatin**.
Orlando V, **Paro** R.
Cell. 1993 Dec 17;75(6):1187-98.

There is a contrast between a sequence-specific DNA-binding protein and the DNA packaging nucleosome.

Histone amino acid side chains do not enter the major groove and 'read' the DNA sequence.



- … but nucleosomes are very crowded along the DNA. Tight wrapping of DNA in nucleosomes means they have positional preferences to bind over more bendable A-T rich sequences first. Very G-C sequences are covered last or not at all.

Most yeast promoters are relatively "nucleosome free" …many mammalian promoters also….

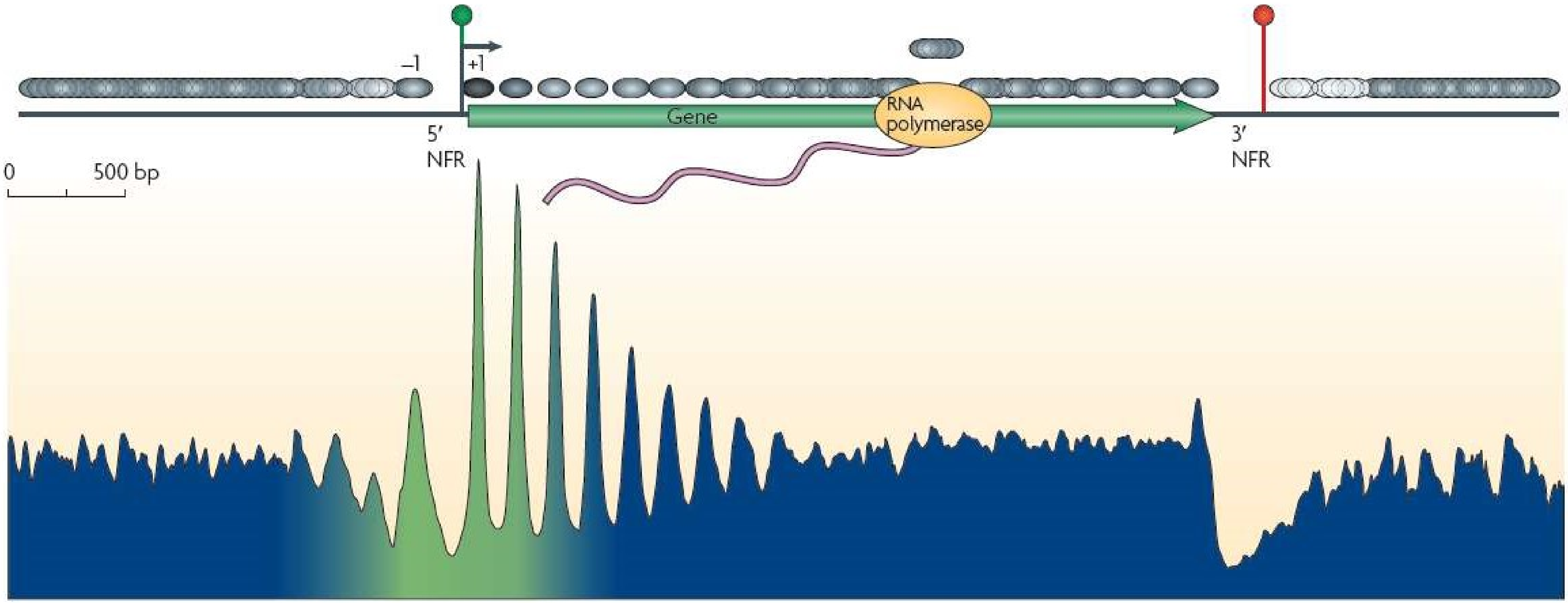


Figure 2 | **Nucleosomal landscape of yeast genes.** The consensus distribution of nucleosomes (grey ovals) around all yeast genes is shown, aligned by the beginning and end of every gene. The resulting two plots were fused in the genic region. The peaks and valleys represent similar positioning relative to the transcription start site (TSS). The arrow under the green circle near the 5′ nucleosome-free region (NFR) represents the TSS. The green –blue shading in the plot represents the transitions observed in nucleosome composition and phasing (green represents high H2A.Z levels, acetylation, H3K4 methylation and phasing, whereas blue represents low levels of these modifications). The red circle indicates transcriptional termination within the 3′ NFR. Figure is reproduced, with permission, from REF. 20 © (2008) Cold Spring Harbor Laboratory Press.

Relatively nucleosome-free regions, DNAse hypersensitive sites and specific histone modifications help identify enhancers in chromatin.

- Chromatin immunoprecipitation methods are easier to use with abundant nucleosomes than with scarce sequence-specific DNA-binding proteins.

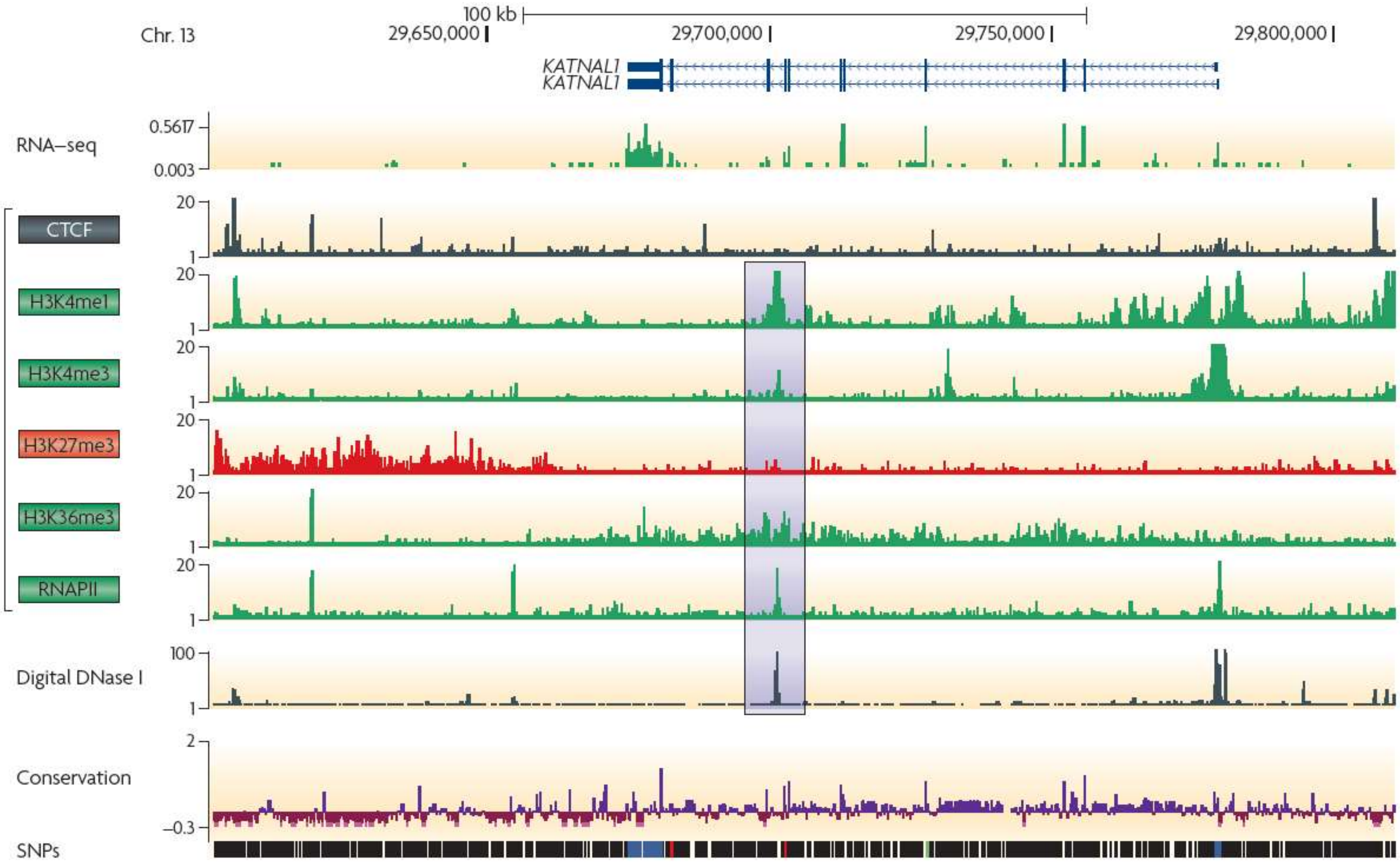
- Modified histones associated with active (H3K4 Me1,Me3) and inactive (H3K27 Me3) genes have been mapped extensively. This is a proxy for mapping the regulatory proteins themselves and allows likely enhancers and promoters to be identified.

- Promoters and enhancers also tend to be nucleosome-free.

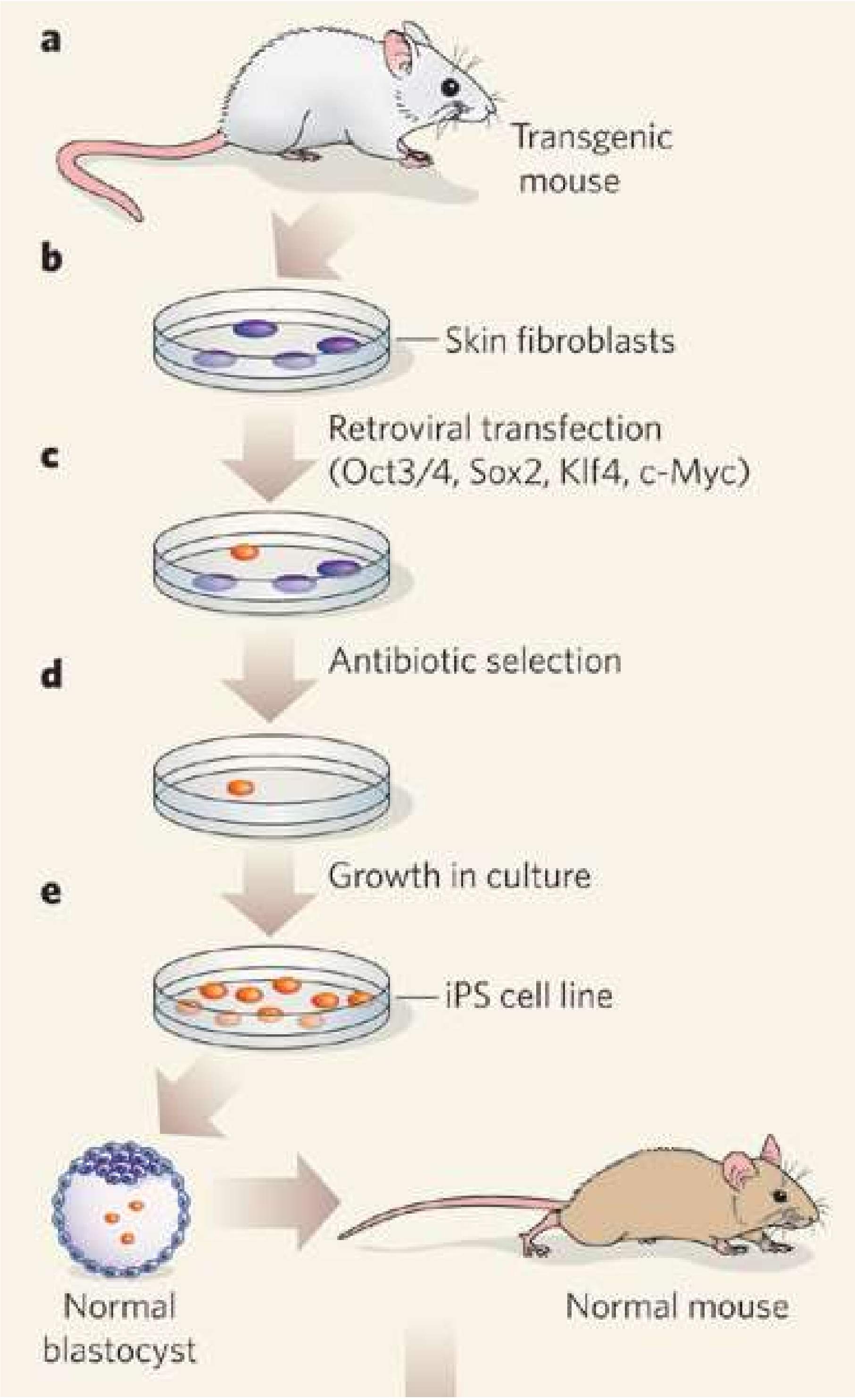
Histone modifications associated with ACTIVE promoters, polII and DNAseI sensitivity. (H3K4 Me1,Me3)

Histone modifications associated with a silent region (H3K27 Me3).

- RNA-seq = reads of cDNA sequence showing where transcripts are coming from.
- CTCF = insulator binding protein, blocks effects of distant enhancers, thought to separate regions of gene control
- SNPs = single nucleotide polymorphisms between human individuals ($3X10^6$ per person!). Some of these cause disease – how do we identify which ones?



The Yamanaka experiment.
Four factors for induced pluripotent stem cells (iPS cells)

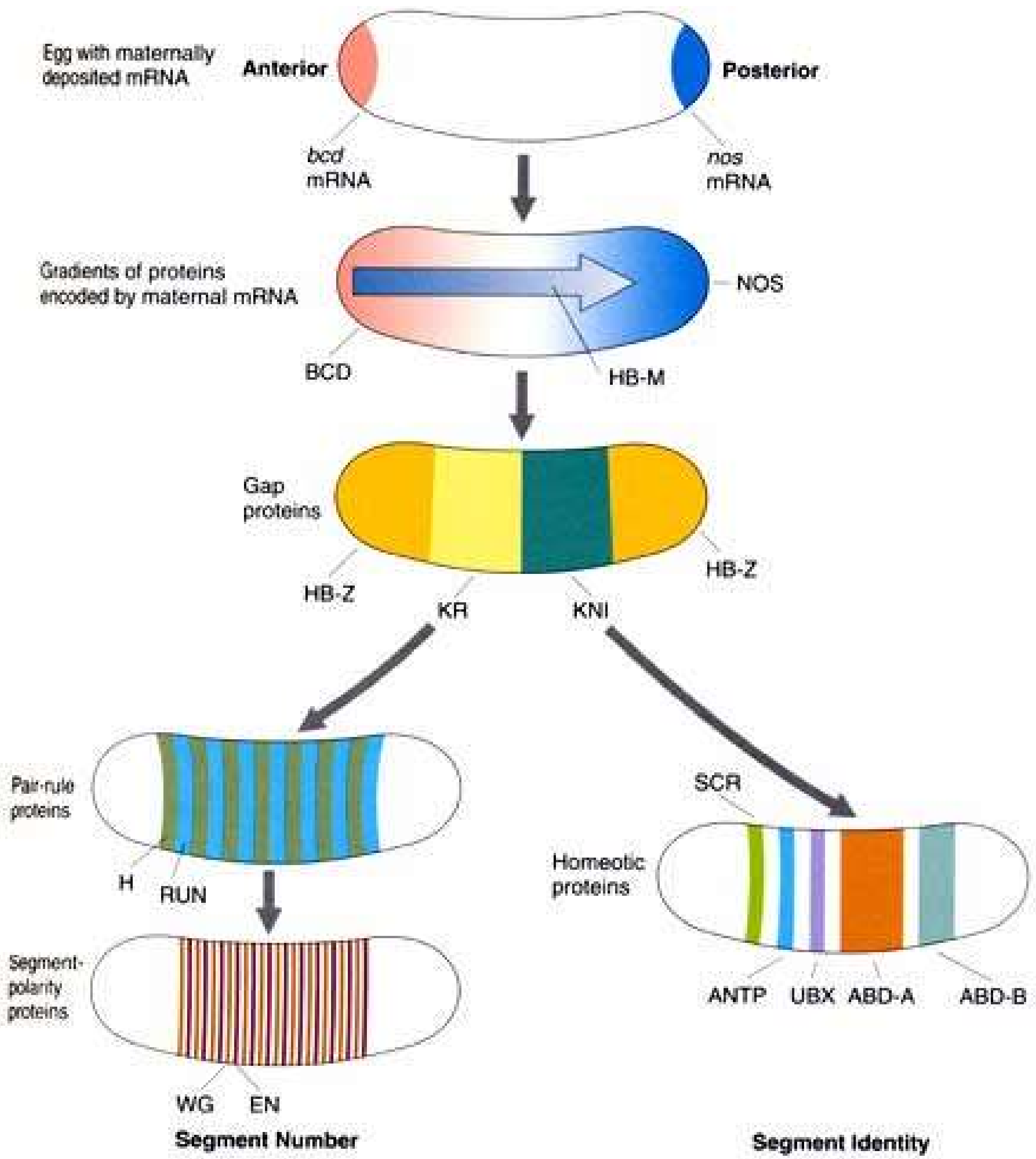


Examples of transcription factors being turned on and acting sequentially and combinatorially

to establish positional coordinates across the developing body

and to control specification of (250?) cell types at the correct positions

A hierarchy of sequence-specific DNA-binding proteins control *Drosophila* embryonic segmentation.

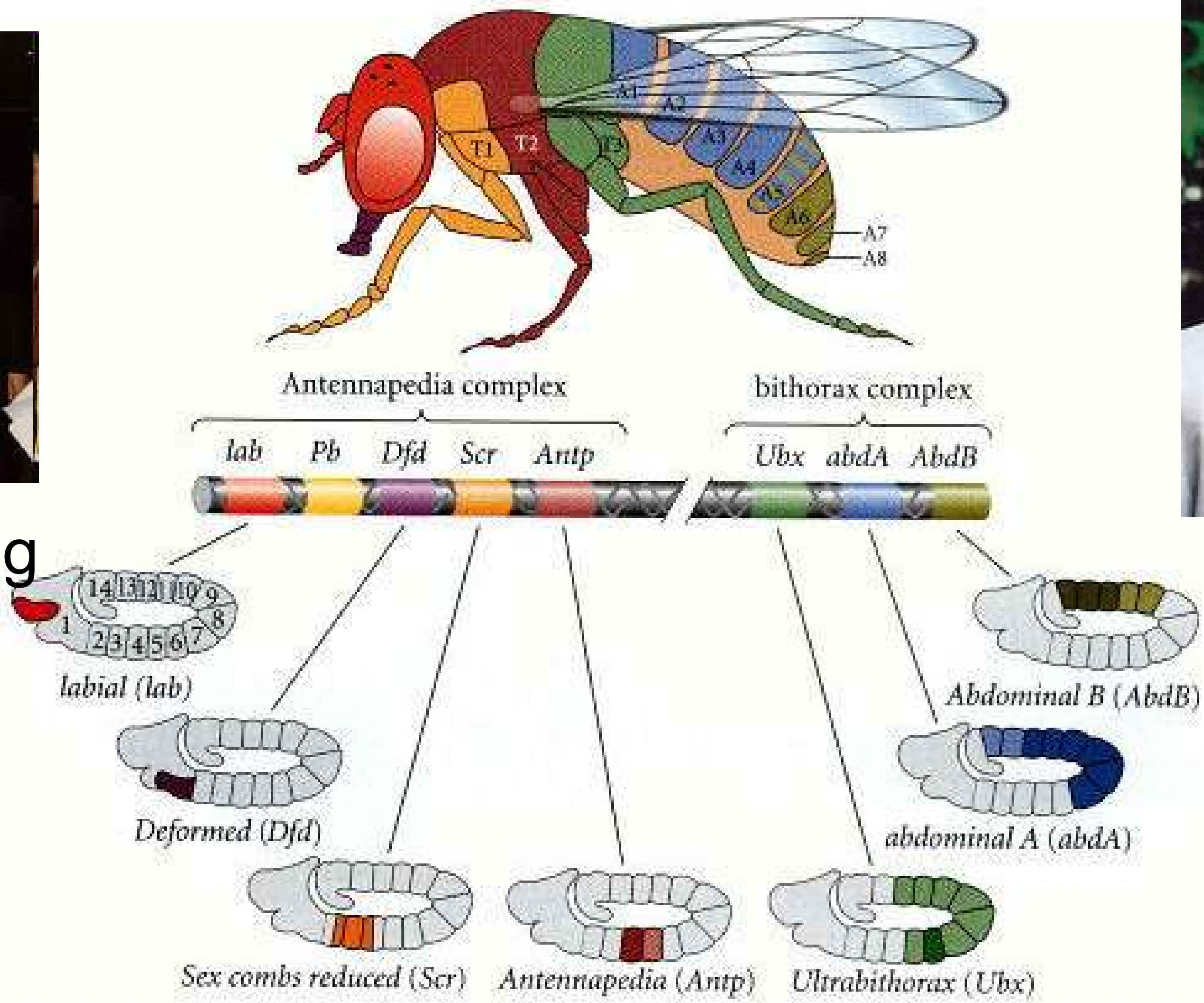


Christiane Nüsslein-Volhard (1942 -)

Eric Wieschaus (1947 -)

After the segment specification stage the segment identities are set by homeotic genes encoding homeodomain transcription factors.

Walter Gehring



Ed Lewis

Gene regulatory networks (GRNs) with different types for feedback and feedforward effects. Pioneered by Eric Davidson for mesoderm specification in sea urchins

Temporal specification by homeodomain transcription factors turned on in sequentially born neuroblasts. (Chris Doe, Richard Mann and others in *Drosophila,* Oliver Hobert in worms.)

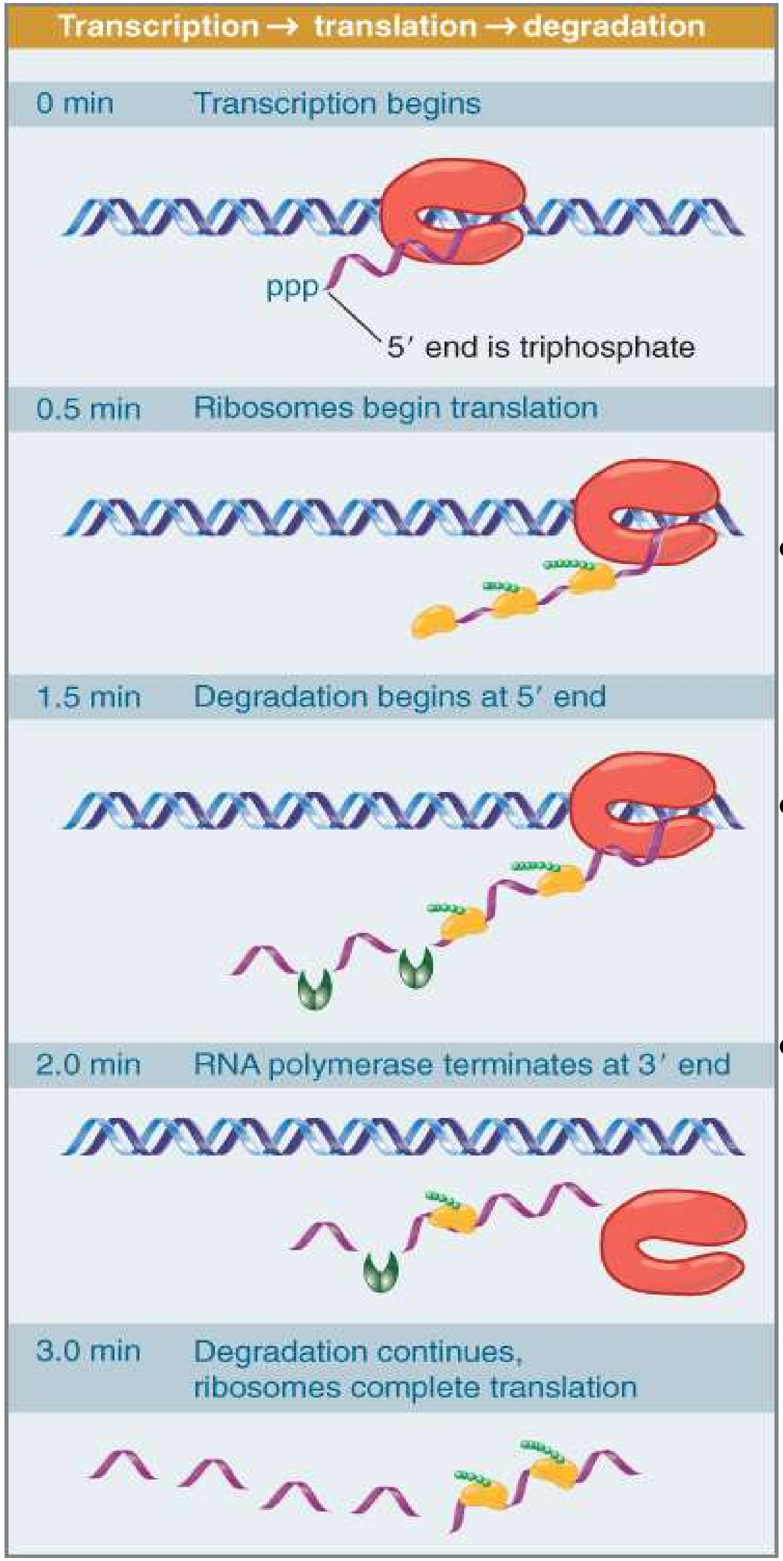
Specification of vertebrate motorneurons by homeodomains (Tom Jessell and Silvia Arber).

Pre-mRNA processing in eukaryotes.
- Ribozymes and the RNA world-

Lecture outline

- Eukaryotic mRNA stability, 5' cap and polyA addition and splicing.
- Mechanism of pre-mRNA splicing and alternative splicing. RNA editing.
- Self-splicing Group II introns, RNA catalysis, Ribozymes and the The RNA World.

Capping and polyA tailing of pre-mRNAs.



The life cycle of a bacterial mRNA.

- Bacterial mRNA is unstable and has a half-life of only a few minutes.
- Transcription in both bacteria and eukaryotes is not very fast at 40 nucleotides/sec, 2.4 kb/min.
- Most bacterial mRNAs can be transcribed in a few minutes whereas a human transcript of 100 kb would take 40 minutes.

Figure 7.14

• hnRNA/pre-mRNA processing

- A typical eukaryotic mRNA:
-
- 5'cap..............___________________________________.....................AAAAAAAAAn
- nontranslated coding long 3' trailer n =100-200
- leader <300nt often >1000nt

Pre-mRNA 5’ cap addition and 3’ polyadenylation and splicing are completed before export from the nucleus.

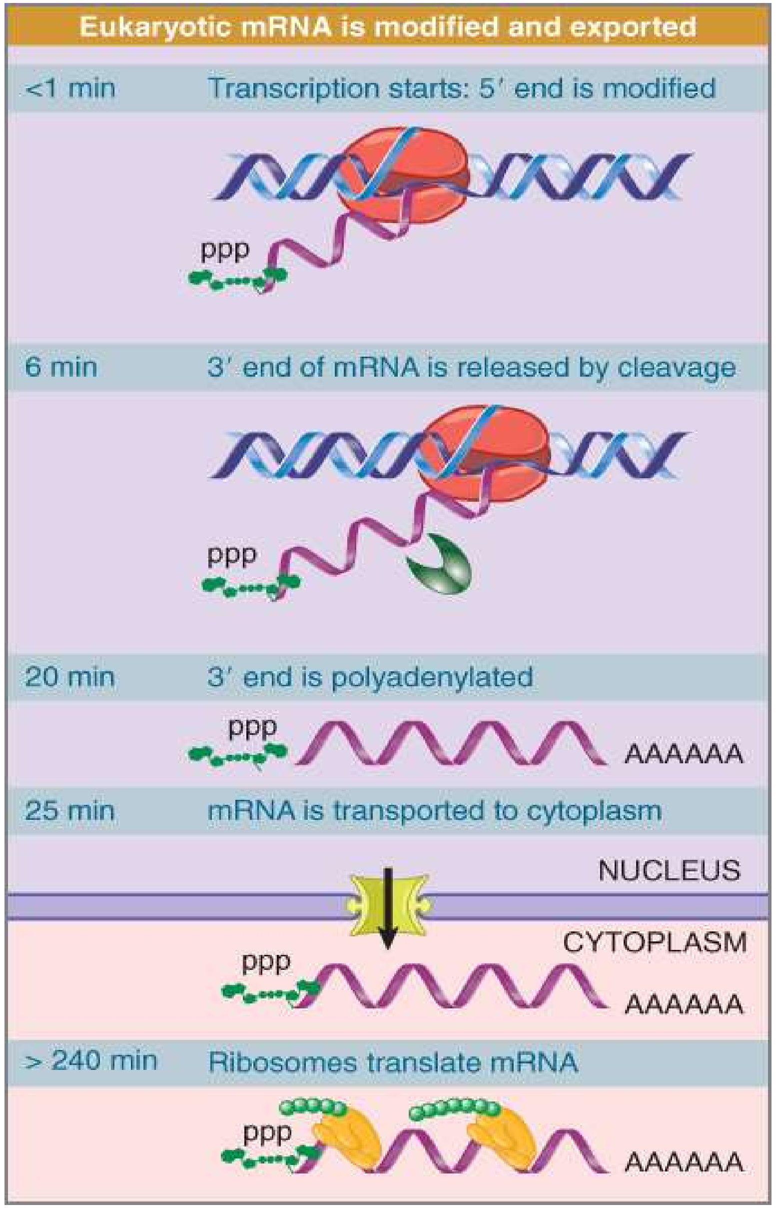


Figure 7.17

The 5′ End of Eukaryotic mRNA Is Capped to protect it against 5’ exonucleases.

- A 5′ cap is formed by adding a G to the terminal base of the transcript via a 5′–5′ link.
- The capping enzyme is **guanylyl transferase.**
  - 1-3 methyl groups are added to the base of the new terminal guanosine
  - 2 methyl-riboses near cap.
  - Base modification m6A in 5’ cap and also inside mRNAs at RRACH consensus, especially around stop codon.

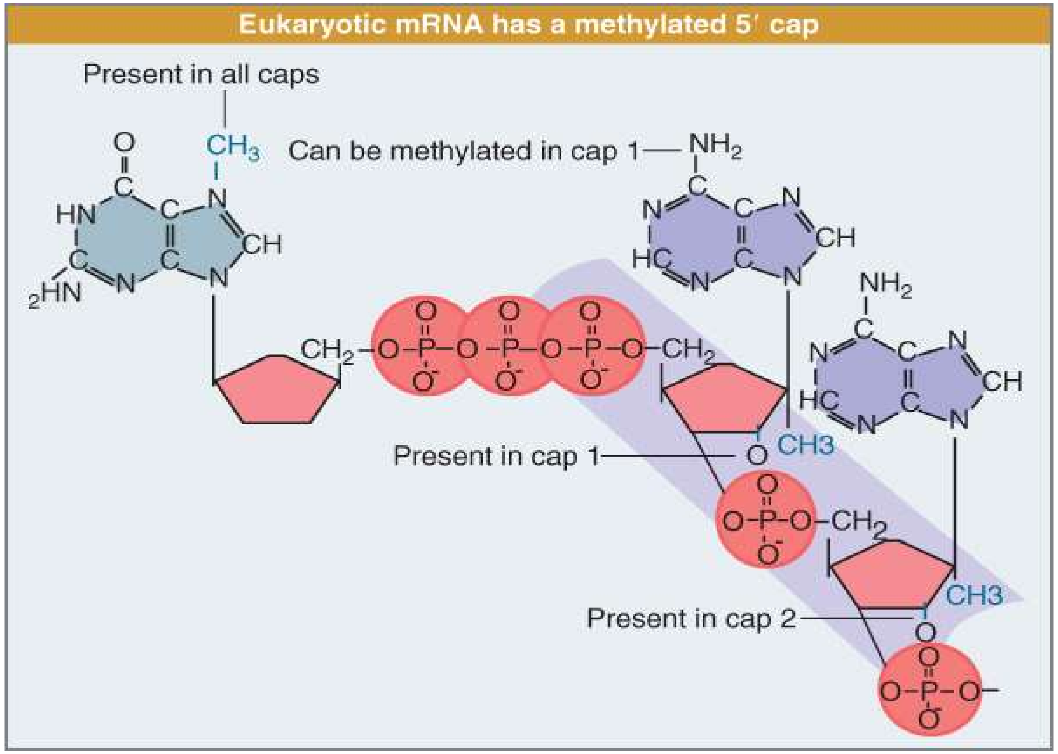
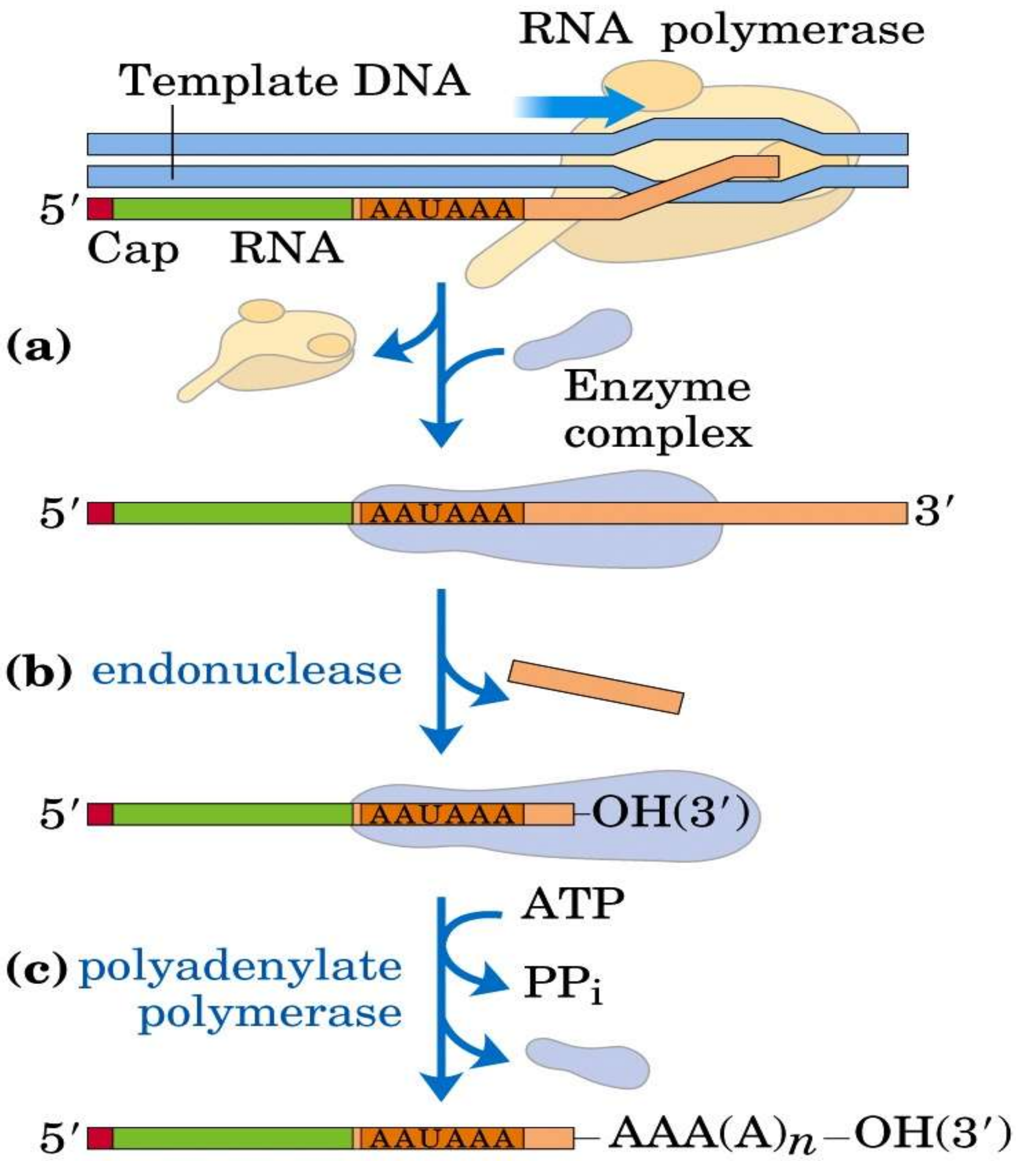


Figure 7.18

3’ end cleavage and polyadenylation.



The cleavage and polyadenylation specificity factor (CPSF) recognizes **AAUAAA (AtwoU, Athree)**.

Cleavage factors (CFI and CFII), endonuclease cleaves RNA 30 bases downstream.

Poly(A) polymerase add ~200 A residues processively to the 3′ end.

The polyA tail stabilizes mRNA, facilitates translation.

- Poly A tail is >200 nucleotides long and binds many copies of Poly A-binding protein (PABP).

- PABP stabilizes the mRNA against degradation.

- PABP binds eIF4G (eukaryotic initiation factor 4G) and facilitates translation of message by the ribosome.

- 5’ and 3’ ends are thought to contact one another by forming a loop of the mRNA.

Removal of cap and polyA tails are required for regulated degradation of mRNA.

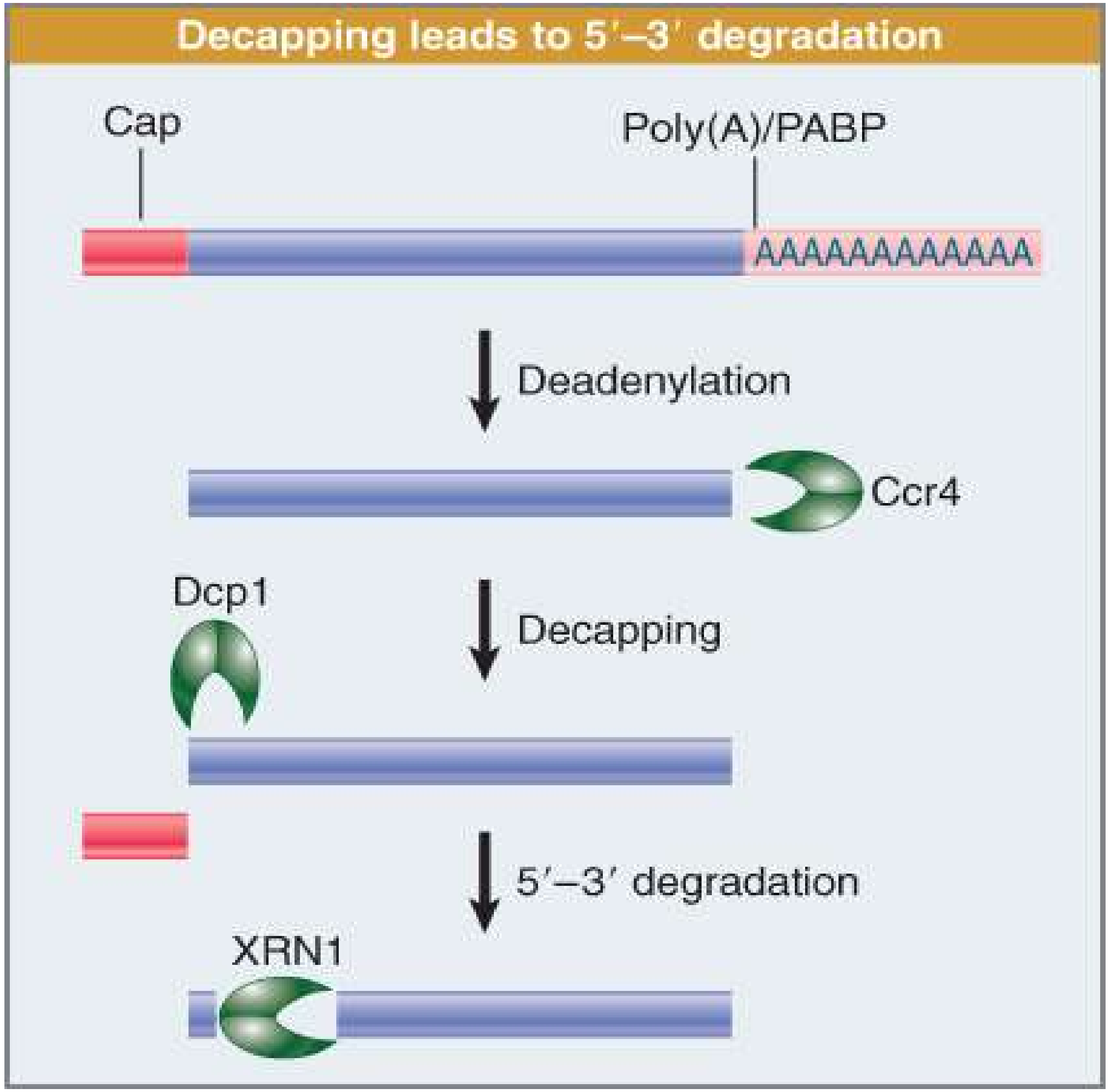


Figure 7.24

- Degradation of yeast mRNA requires removal of the 5′ cap and the 3′ poly(A).
- One yeast pathway involves exonucleolytic degradation from 5′→3′.

- Another yeast mRNA degradation pathway uses a complex of several exonucleases that work in the 3′→5′ direction The **exosome**.

- The deadenylase of animal cells may bind directly to the 5′ cap.

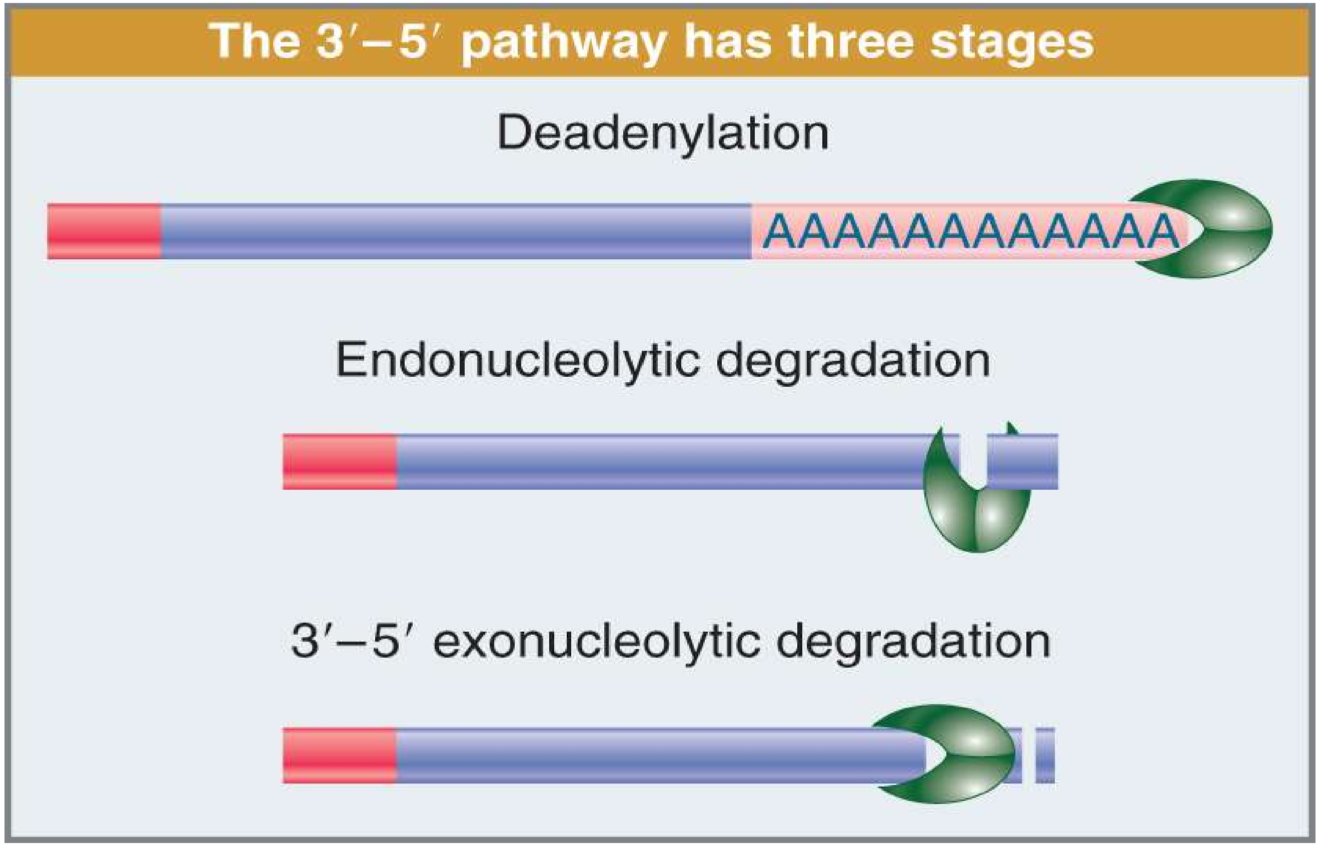


Figure 7.25

The mechanism of eukaryotic pre-mRNA splicing.

Pre-mRNA splicing.

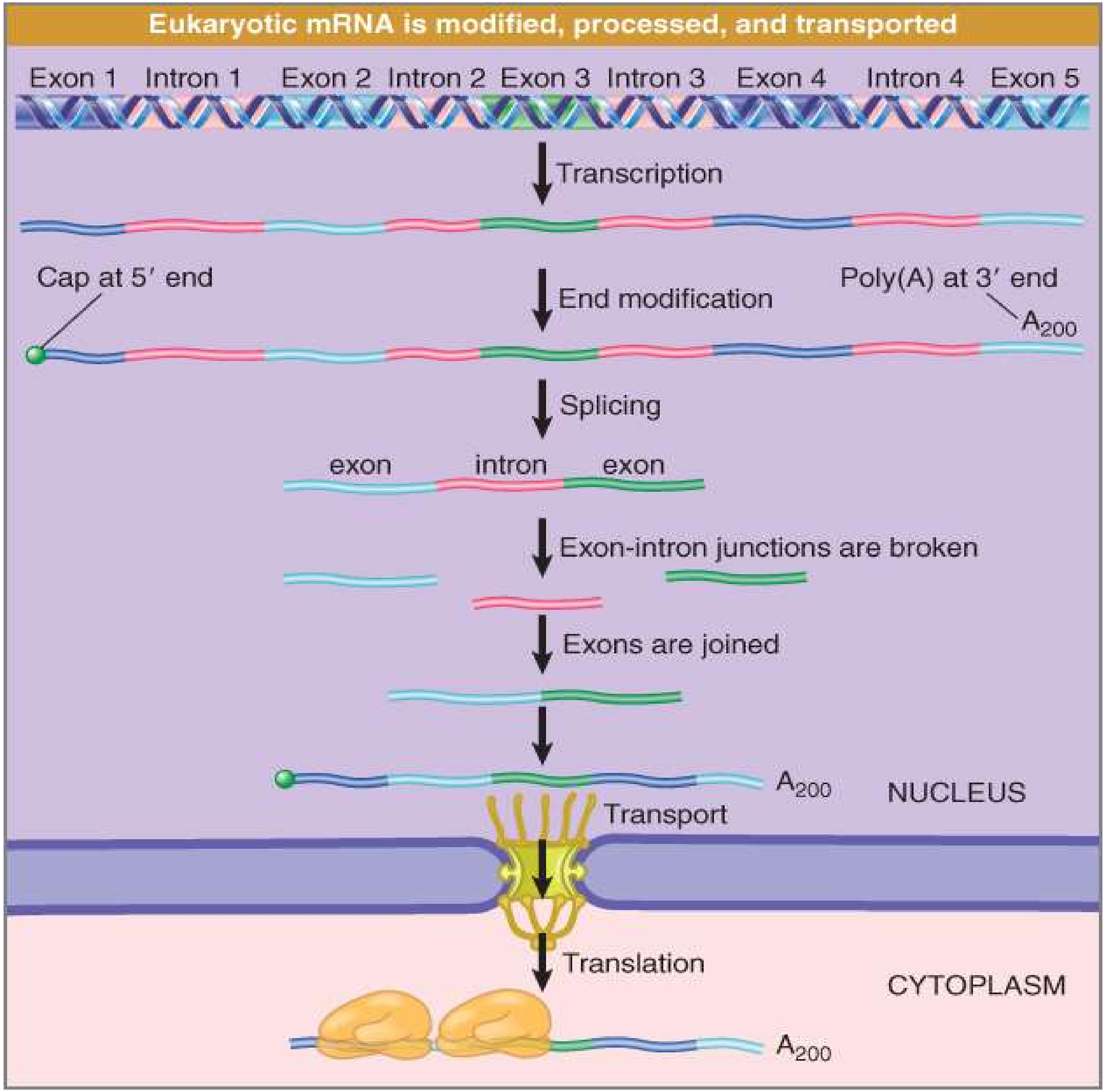
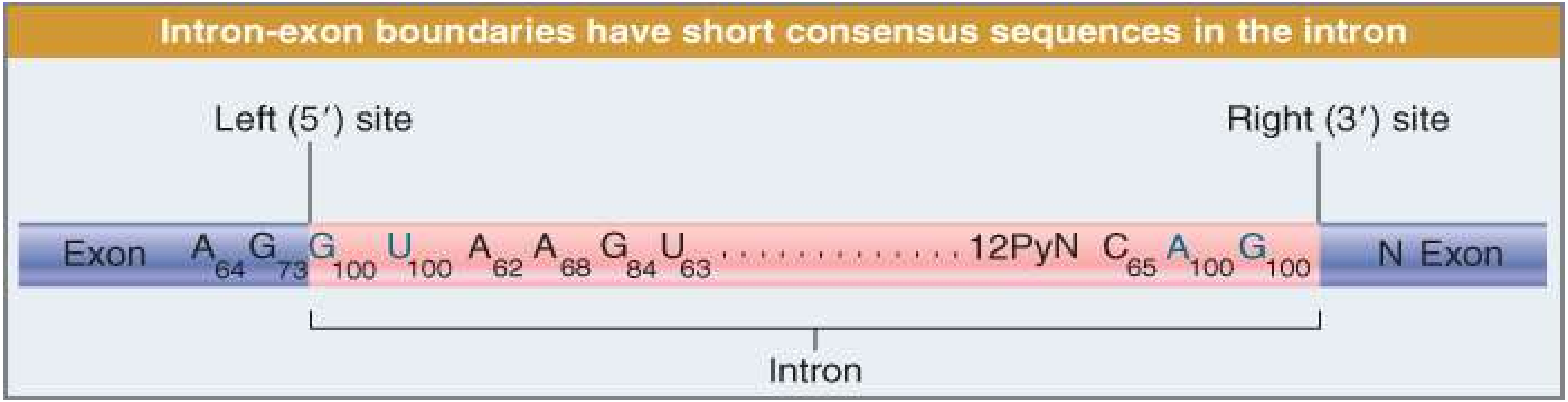


Figure 26.2

Consensus 5' and 3' splice sites.
The GU-AG rule



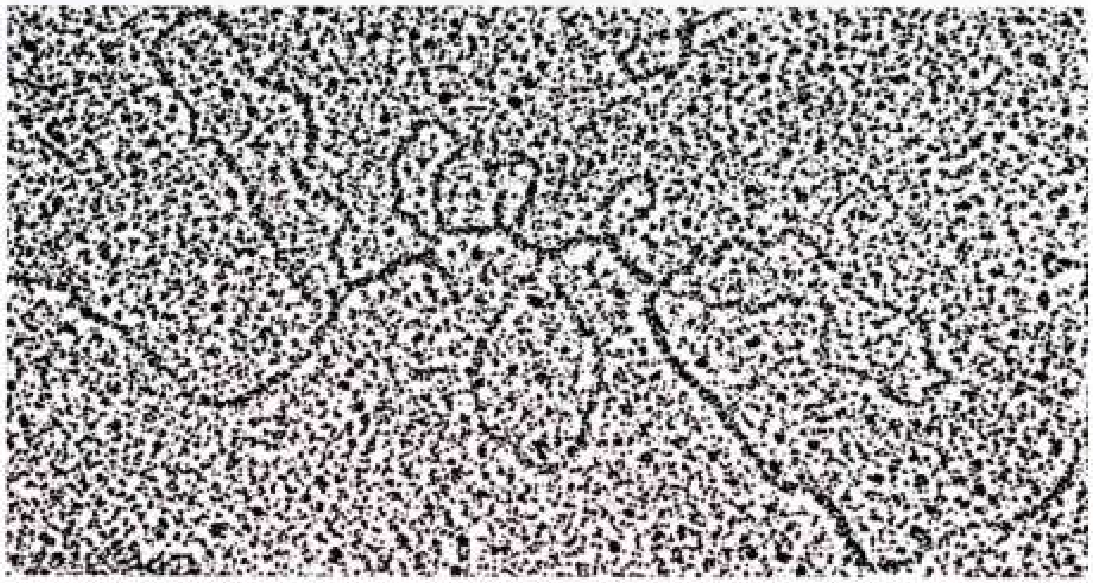
Protein coding sequence must be 'in frame' across the splice junction.

Remember GU…AG are splice junction sequences. Mnemonic; GULAG maybe?

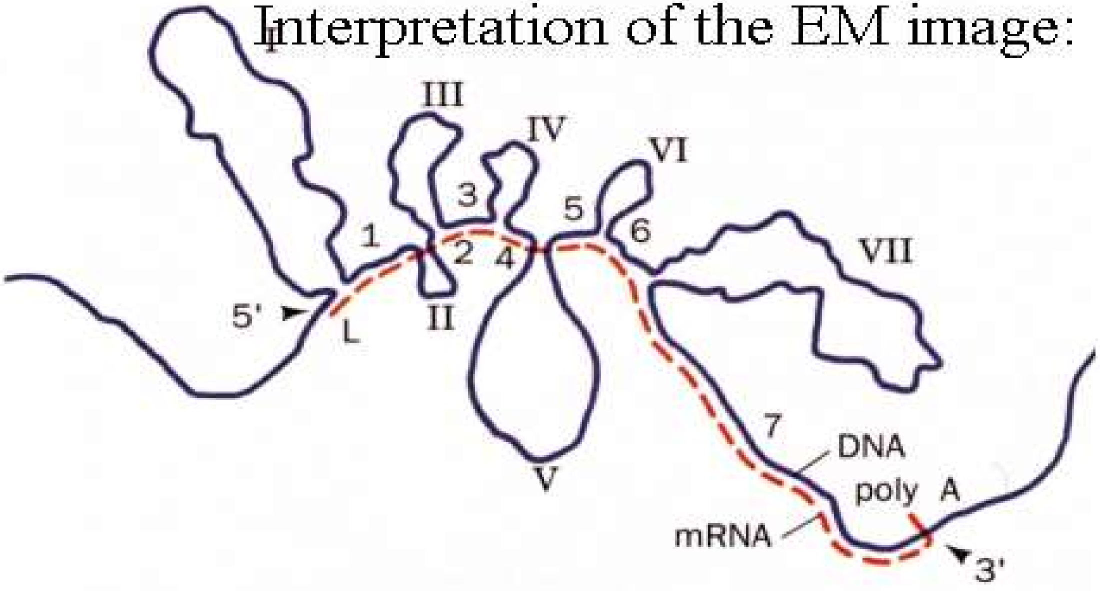
Intron sequences are generally not conserved even between closely related genes.

Figure 26.3

EM heteroduplex analysis of Adenovirus RNA transcripts annealed to Adenovirus DNA showed formation of D (displacement) loops.



Interpretation of the EM image:



Annealing RNA from virus-infected cells with viral DNA revealed the existence of seven introns-transcribed regions of the DNA removed from the mature mRNA.

Splicing must preserve the open reading frame.

- 5’ and 3’ splice site have to be in the same reading frame.
- Reading frames are 1, after first base of codon, 2 after second base and 3 after third base, ie. between codons
- GU..AG consensus are the ends of the intron to be removed, not in the actual coding sequence.

Splice Junctions Are Read in Pairs

- All 5′ splice sites are functionally equivalent, and all 3′ splice sites are functionally equivalent.
- Usually draw exons as boxes to emphasise them with splices joining box corners and introns as lines between the boxes .
- Gene rearrangements within big introns have a good 1/3 chance to produce fusion proteins. (Exon shuffling in evolution).

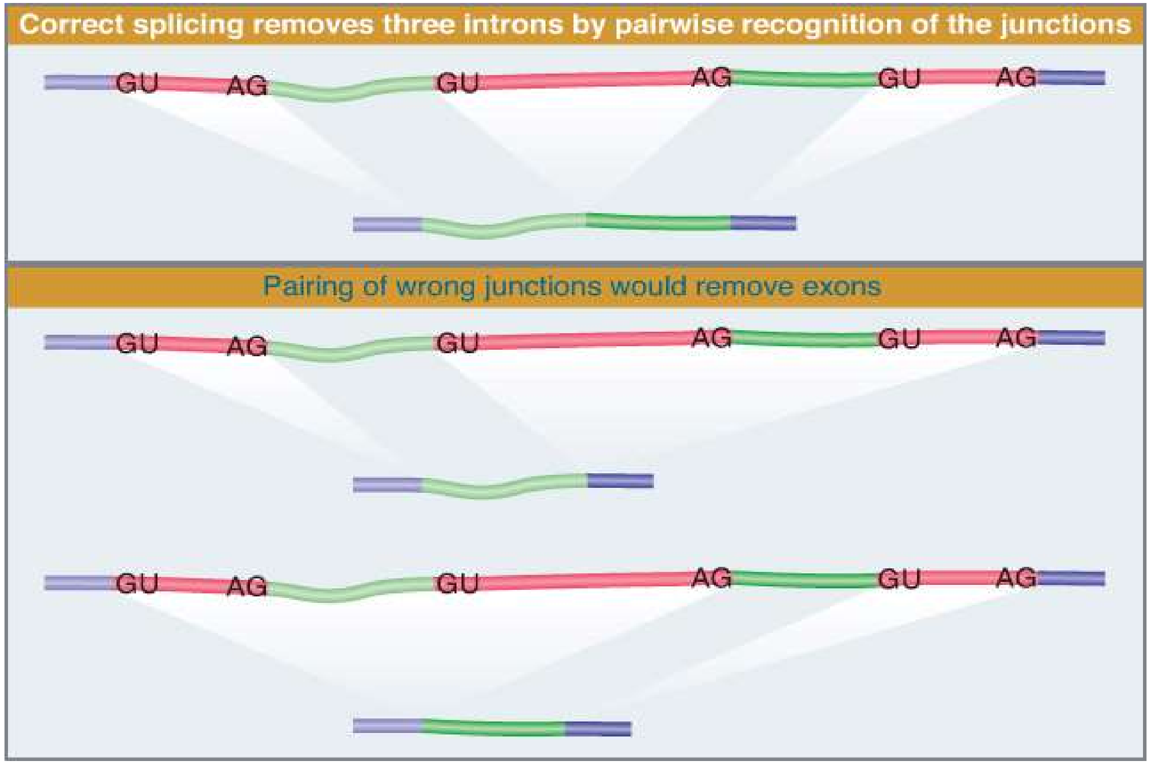
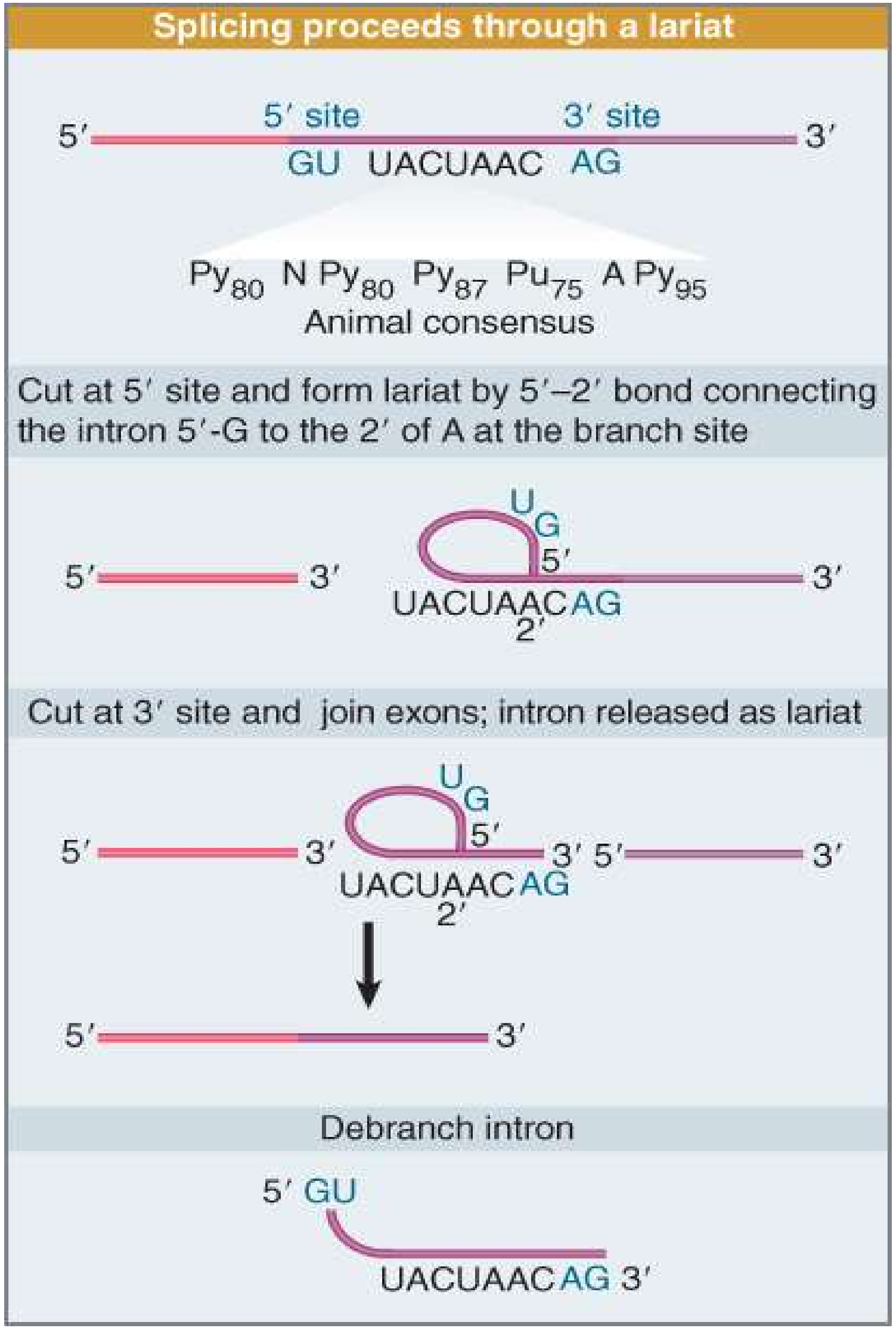


Figure 26.4


Splicing proceeds through a lariat
5′ site
3′ site
5′
3′
GU UACUAAC AG
Py80 N Py80 Py87 Pu75 A Py95
Animal consensus
Cut at 5′ site and form lariat by 5′–2′ bond connecting the intron 5′-G to the 2′ of A at the branch site
UG
5′
3′
UACUAACAG
2′
Cut at 3′ site and join exons; intron released as lariat
Debranch intron
5′ GU
UACUAACAG 3′

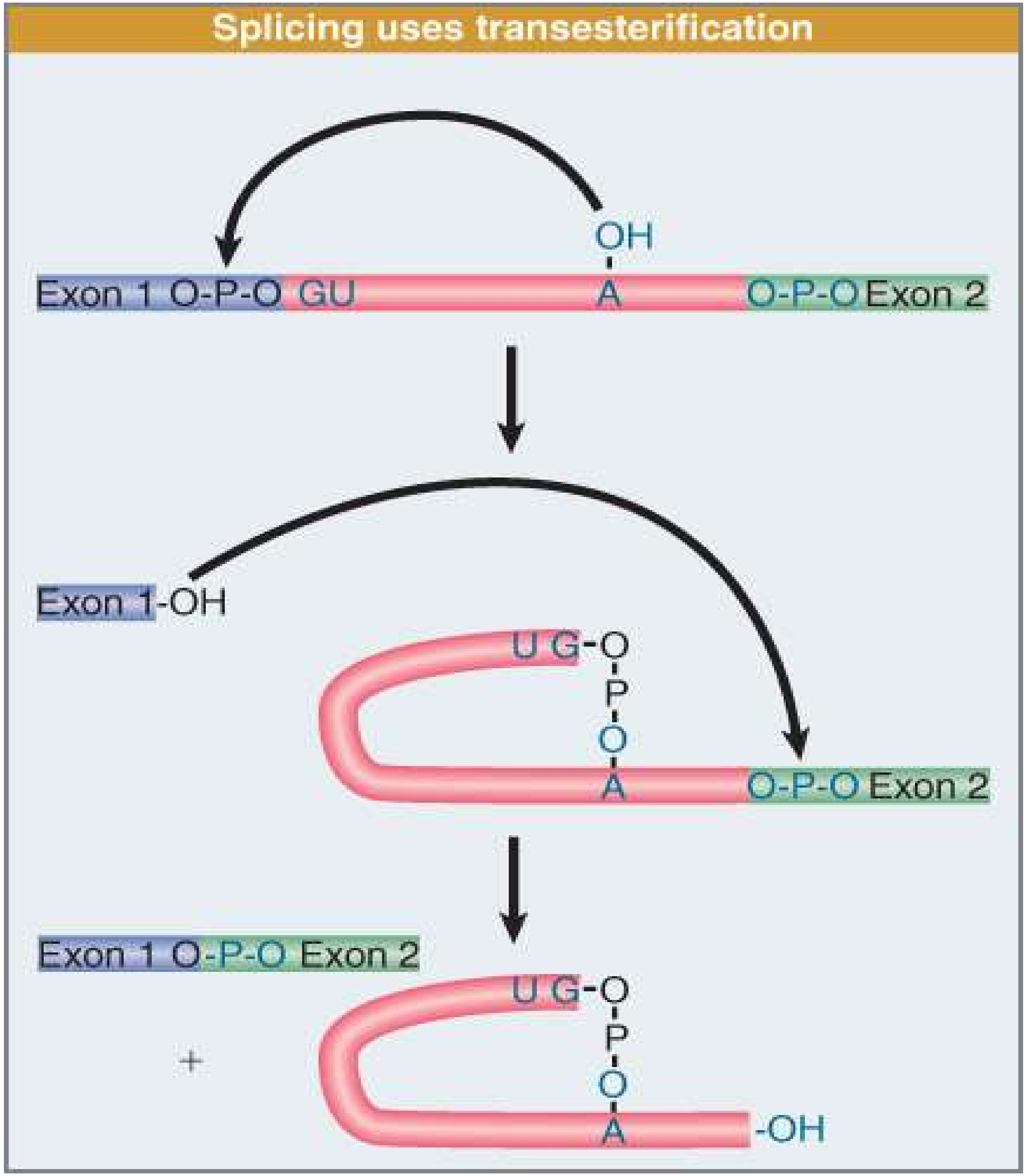


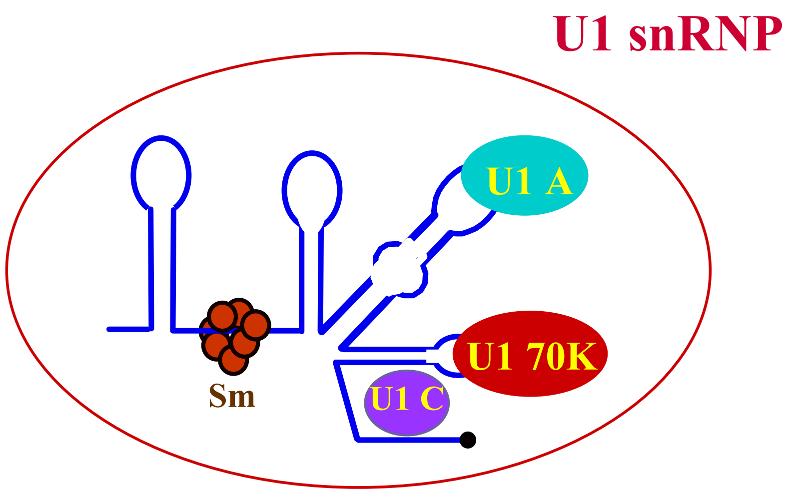
Figure 26.7

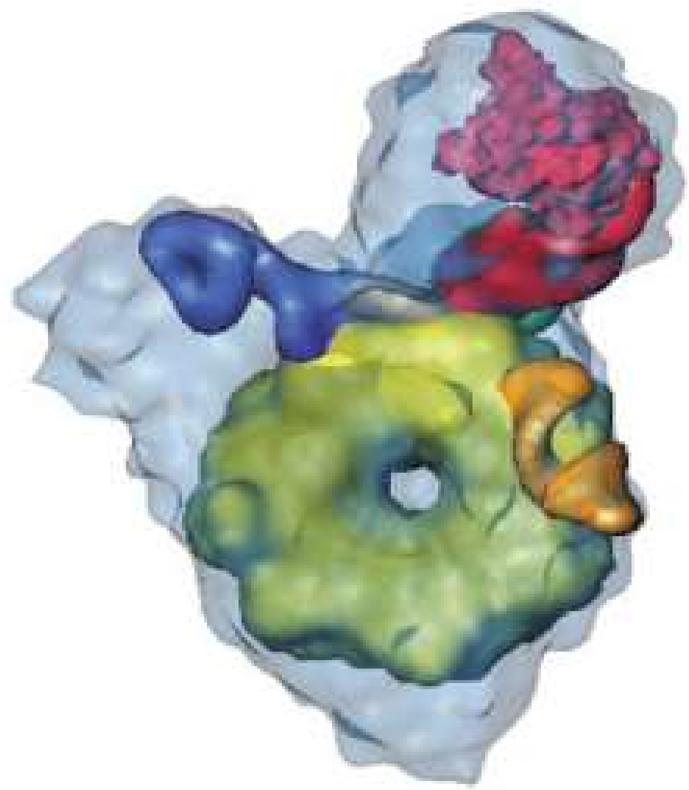
Alkaline lysis of ssRNA.

The reactive 2' OH groups on riboses can also mediate complete cleavage of 5'-3' phosphate bonds in single stranded RNA under alkaline conditions

Splicing is catalysed by a series of five small nuclear ribonucleoprotein particles (snRNPs, 'SNURPs'). Conserved small structured RNAs with Sm and other proteins bound.

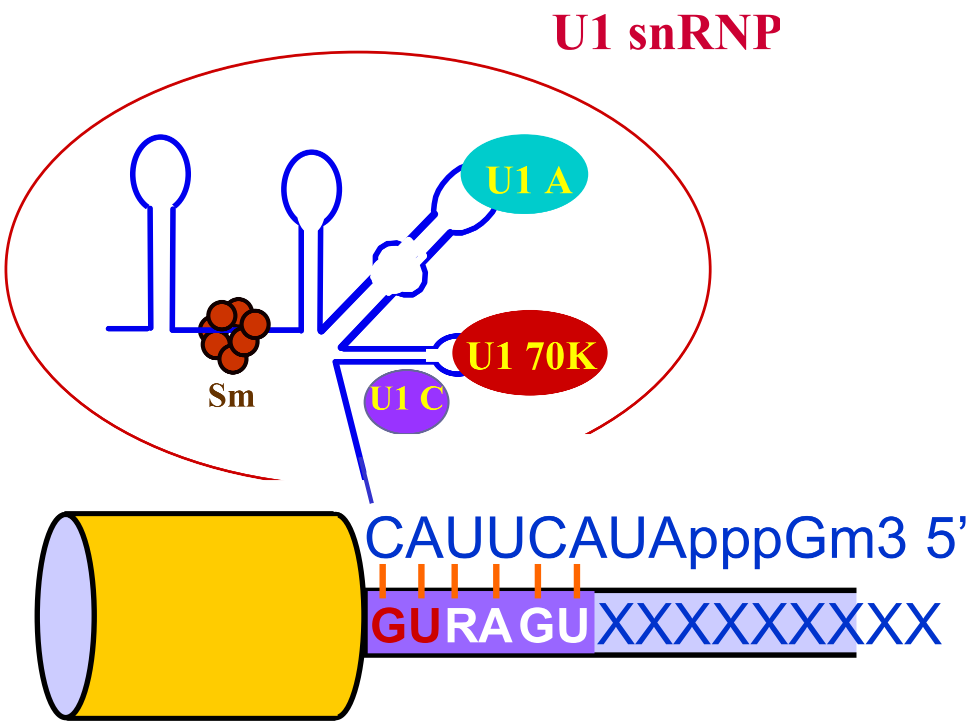
SNURPs assemble sequentially on pre-mRNA to form a <u>spliceosome</u>.



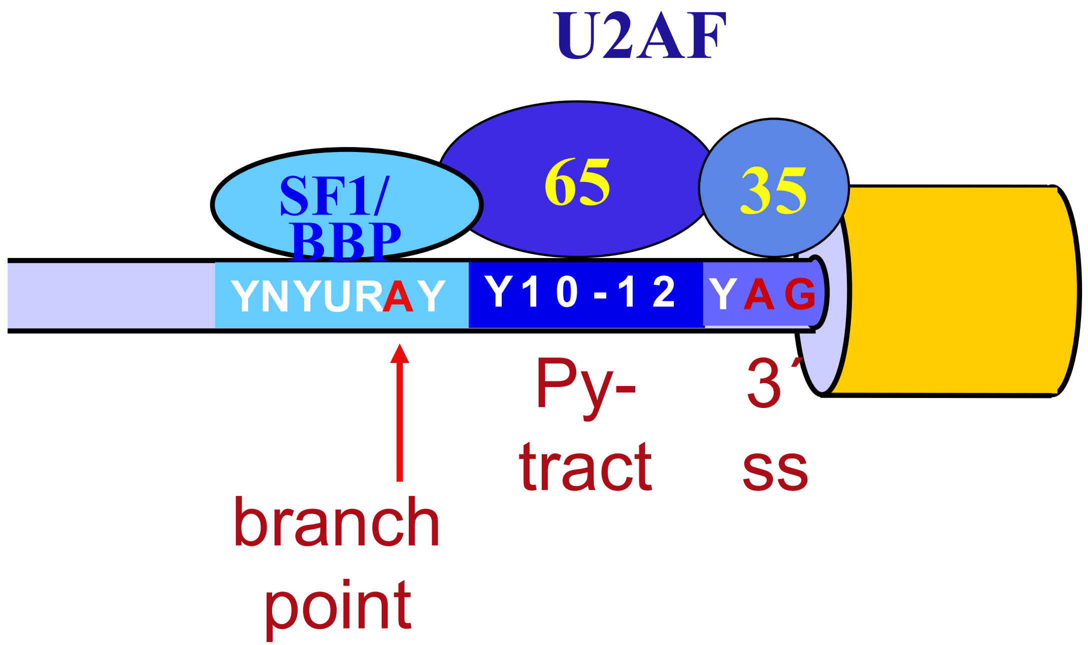


Cryo-EM picture
(Stark et al. Nature 409, 539)

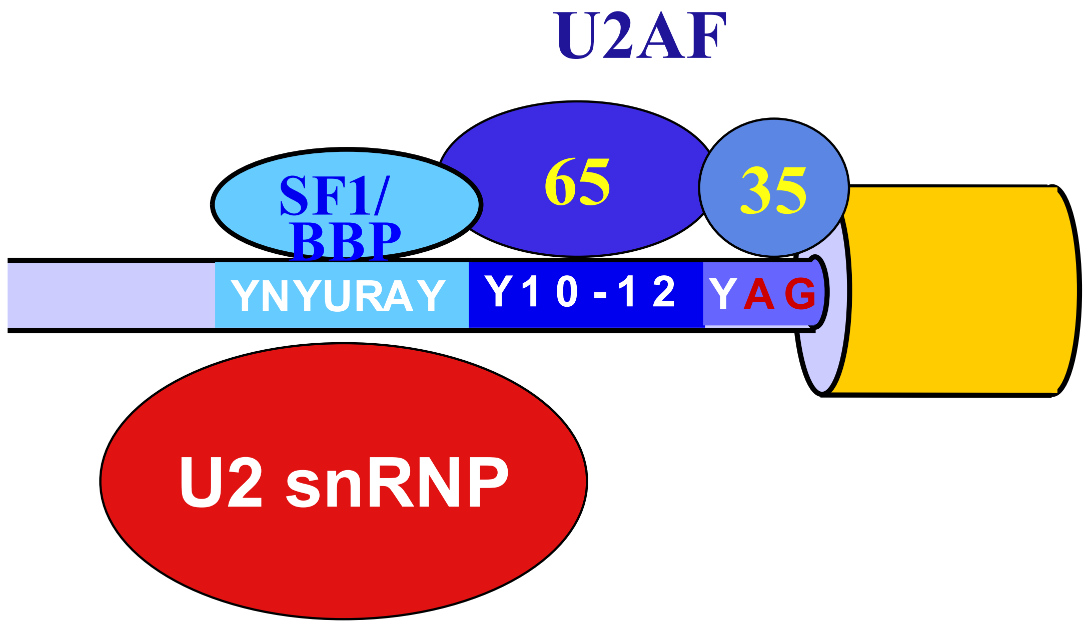
5’ Splice Site Recognition



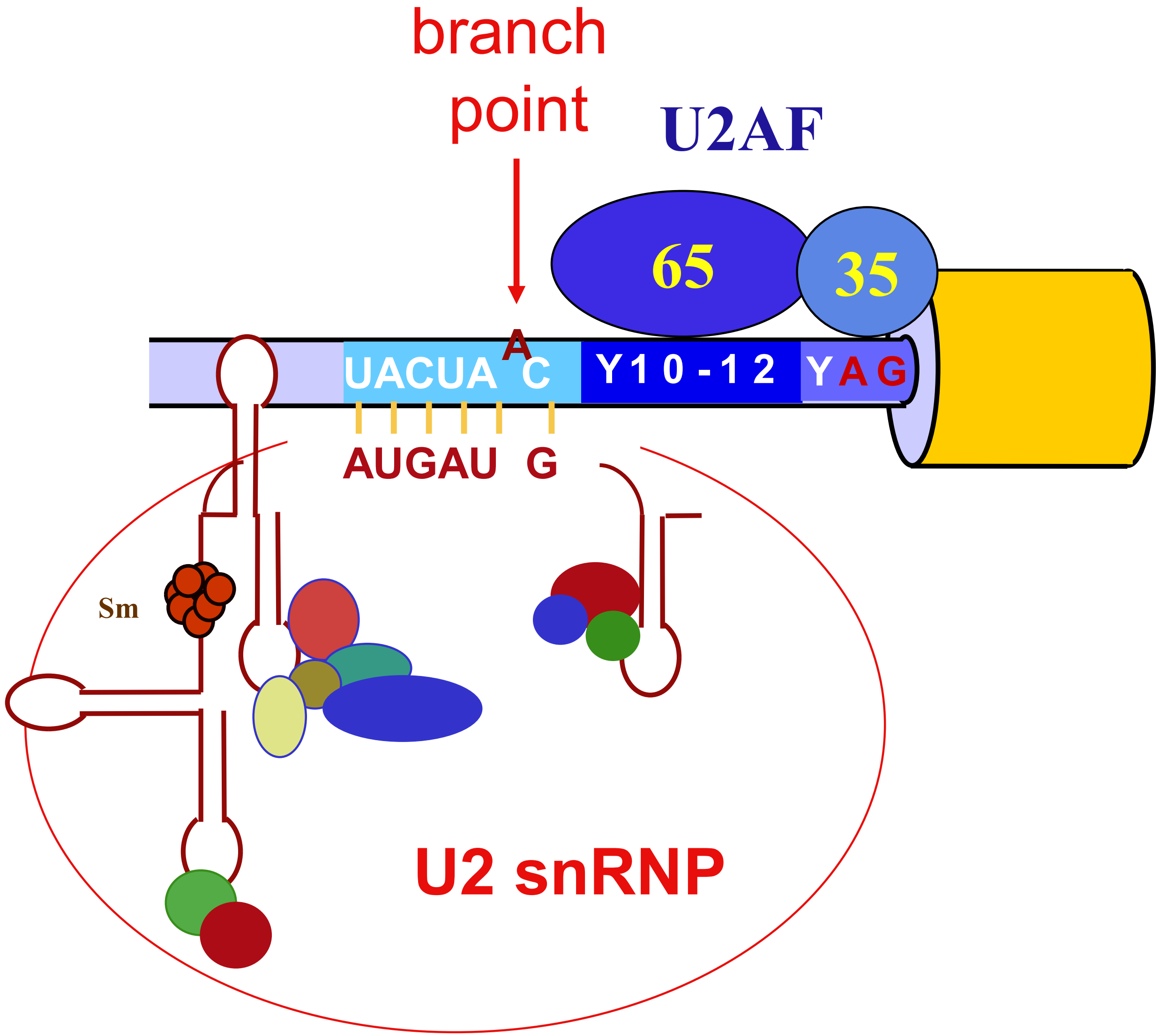
3’ Splice Site Recognition



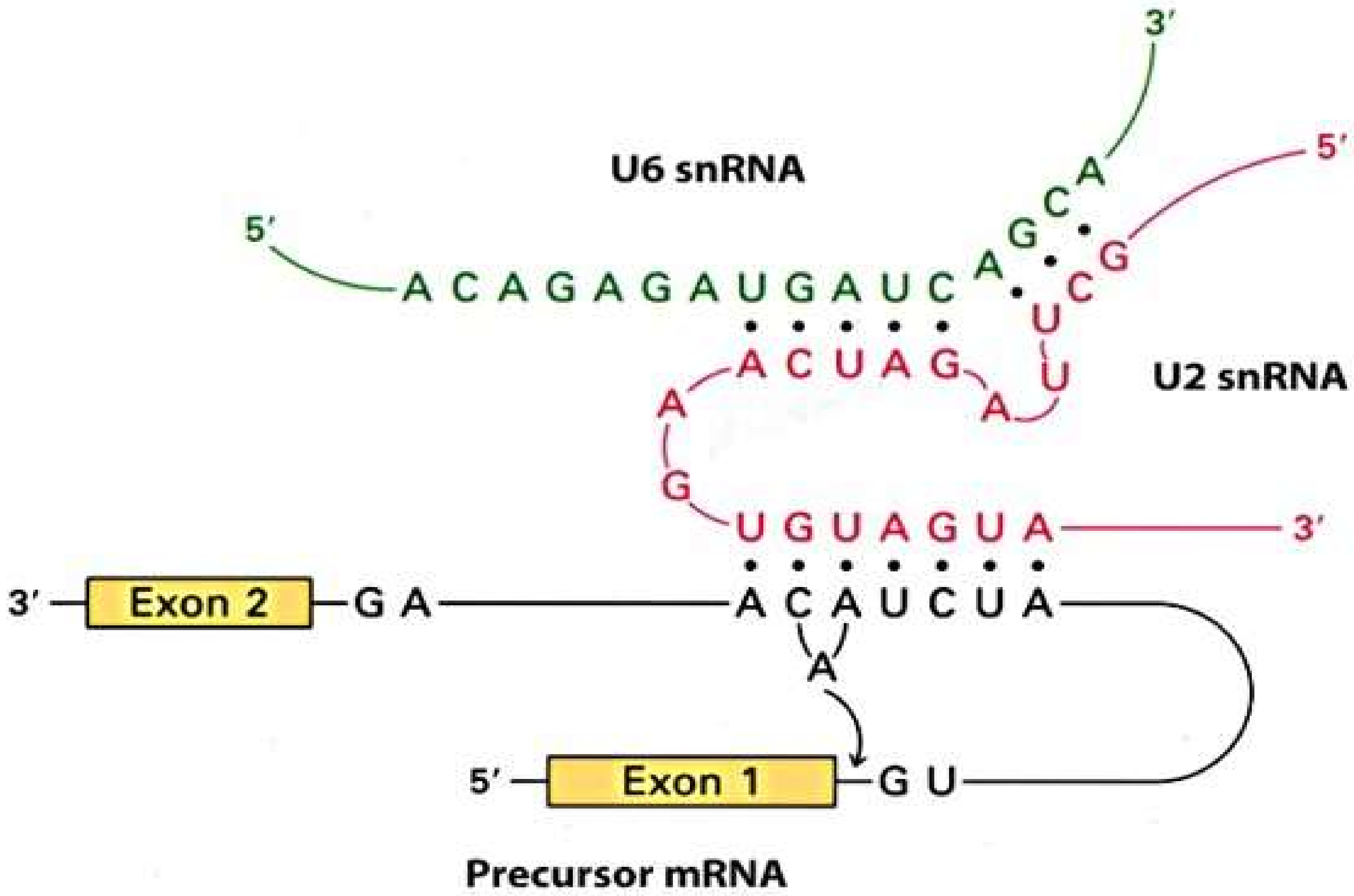
3’ Splice Site Recognition



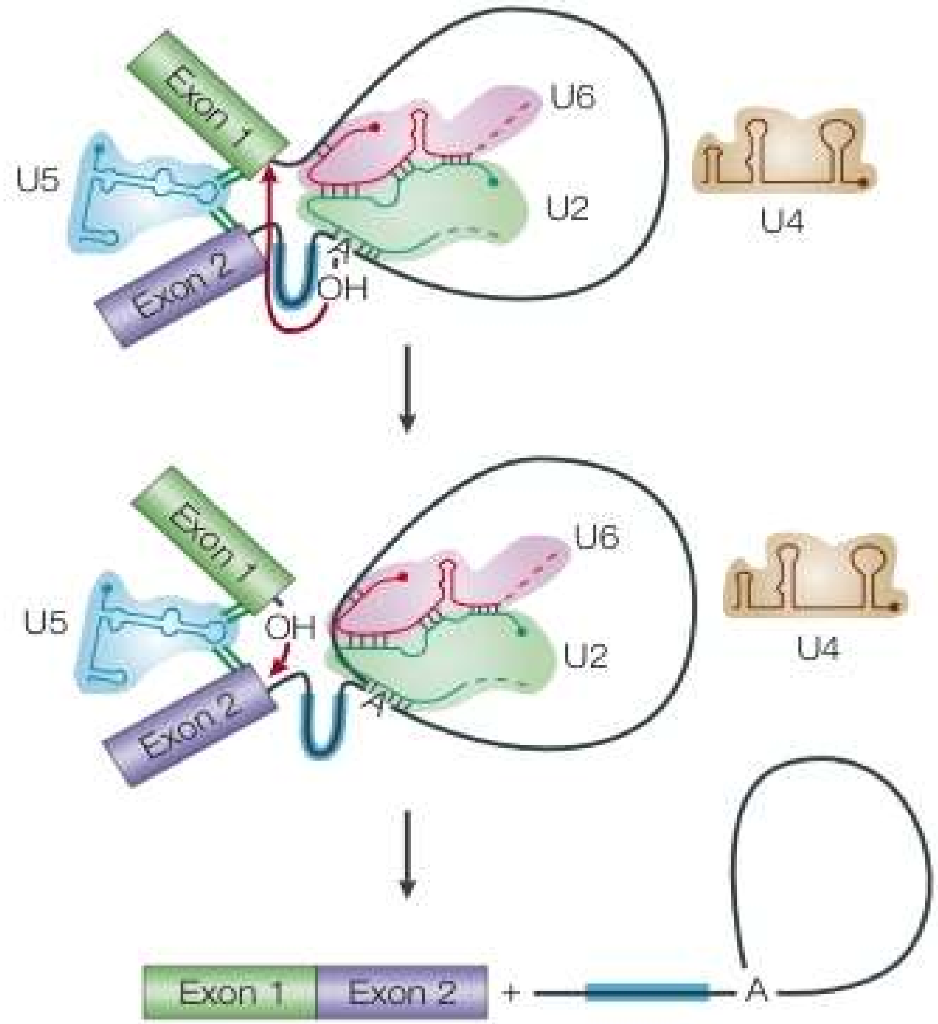
3’ Splice Site Recognition



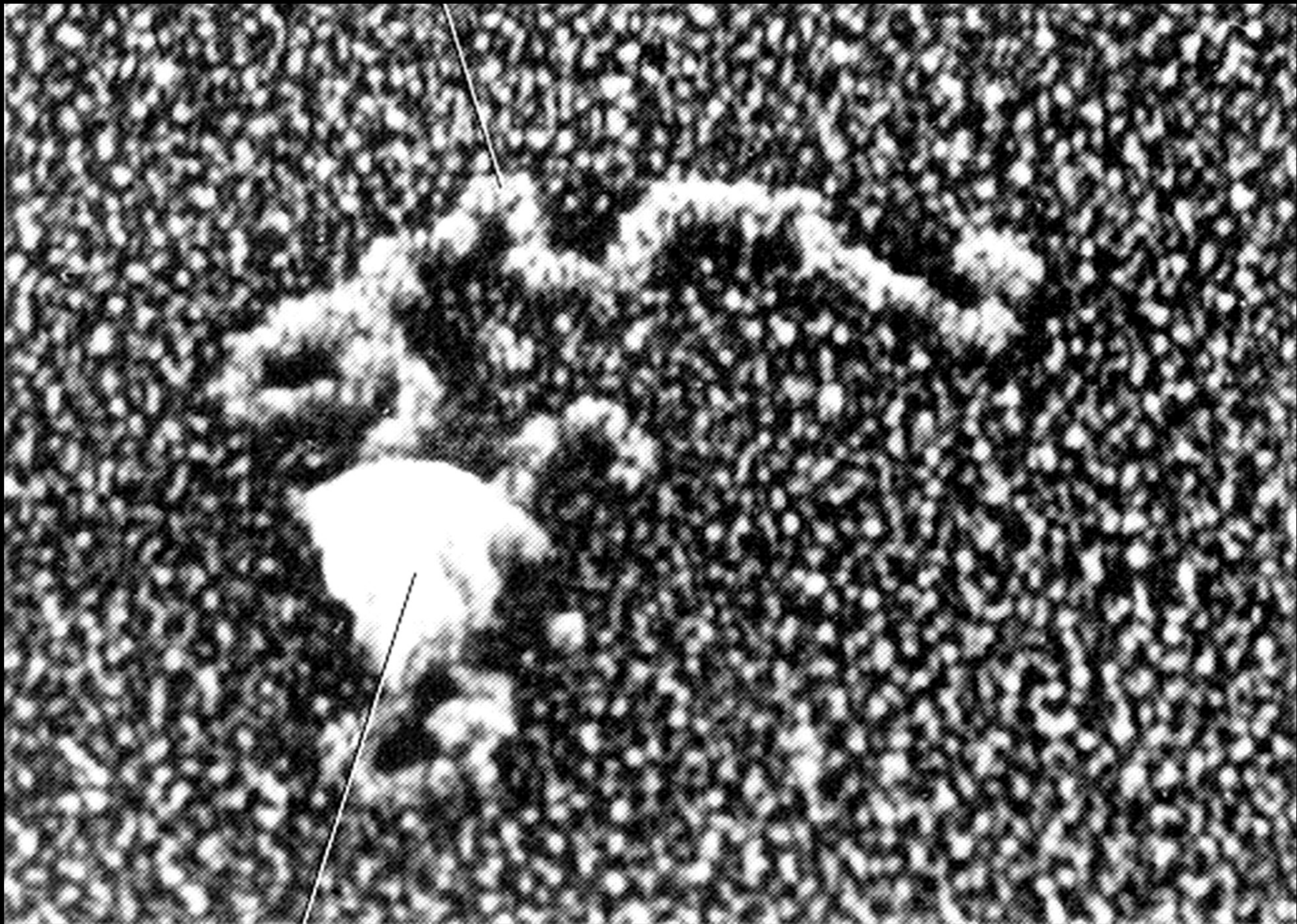
The U4/U5/U6 tri snRNP arrives last and base pairs to U2.



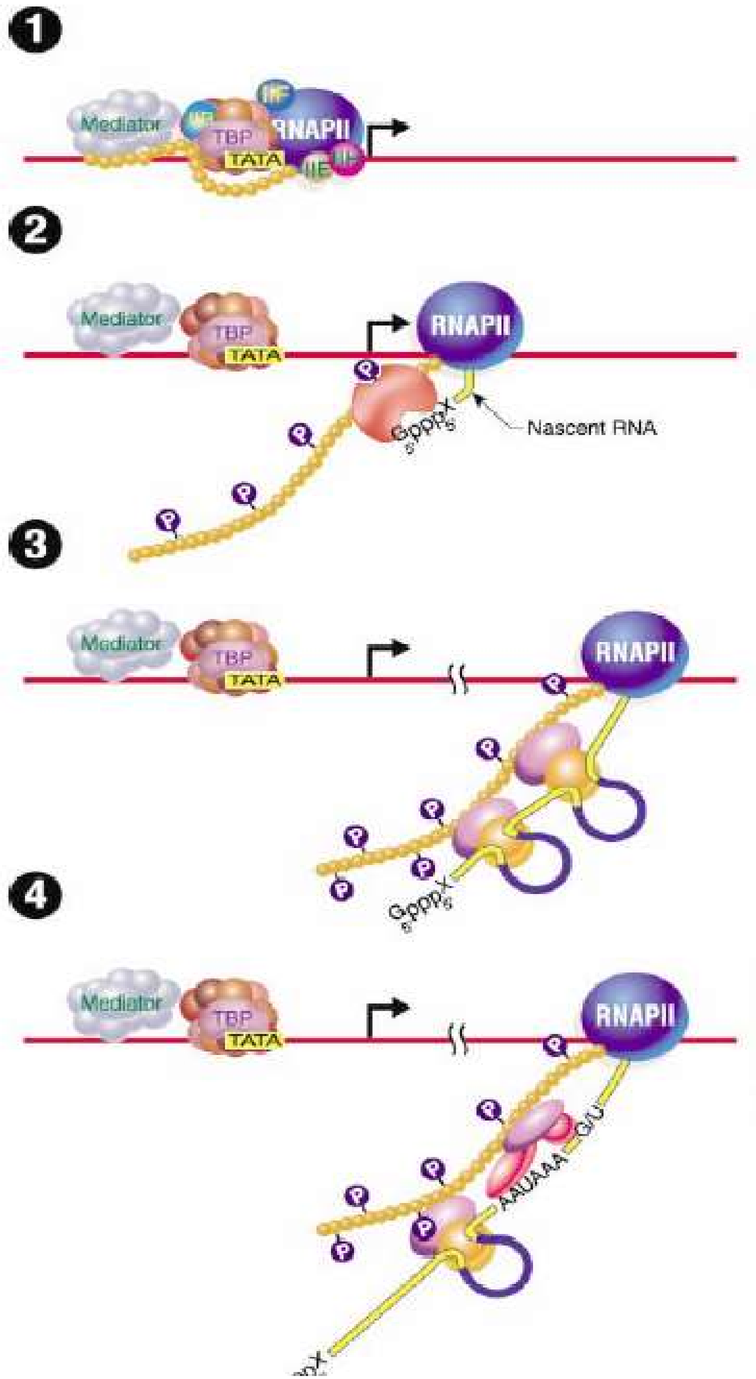
Conformation changes bring splice junctions together.



Patel and Steitz, Nature Reviews Molecular Cell Biology (2001) 4, 960-970


Pre-mRNA
spliceosome

RNA processing is coupled to transcription.



Splicing Is Connected to Export and efficient translation of mRNA

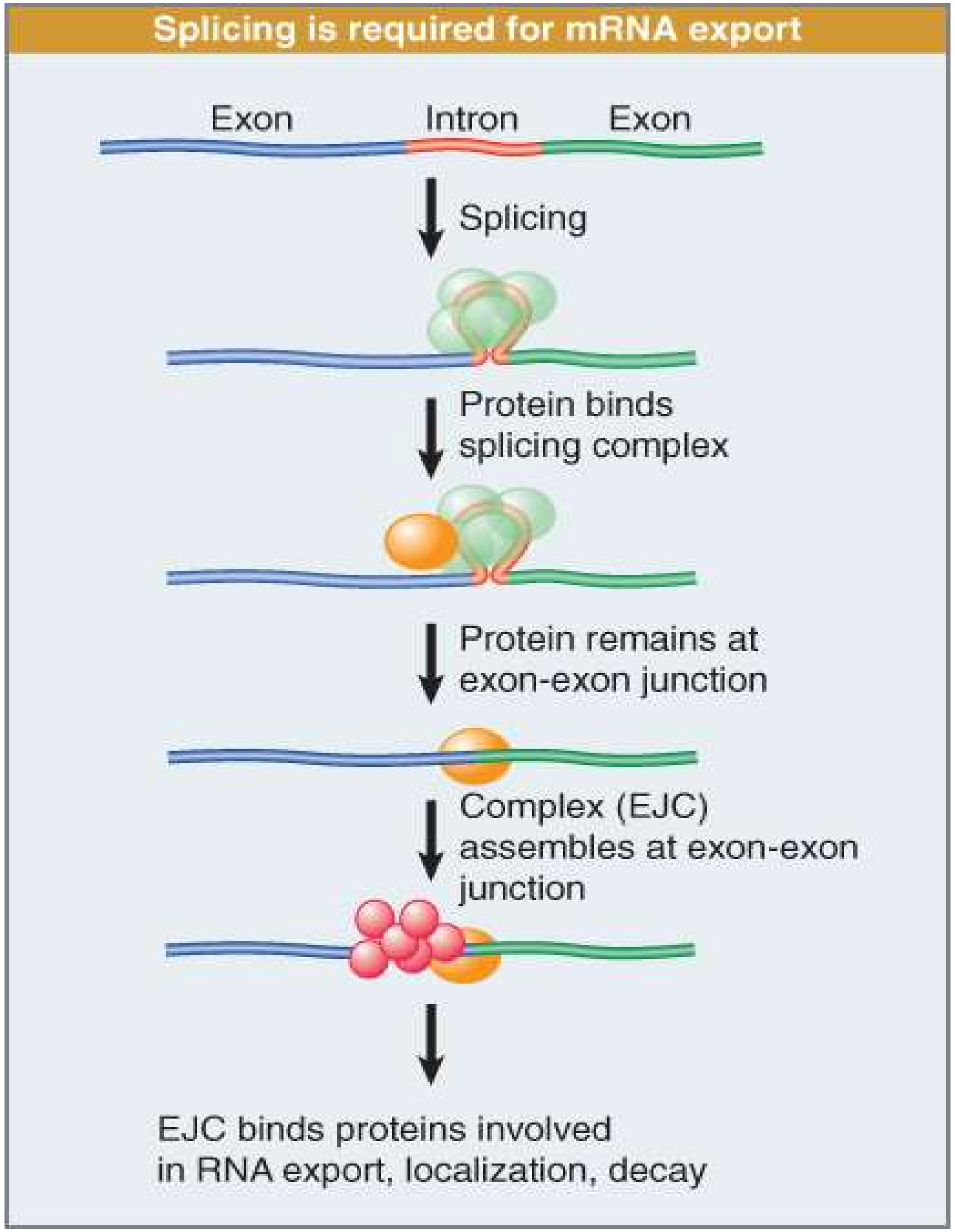
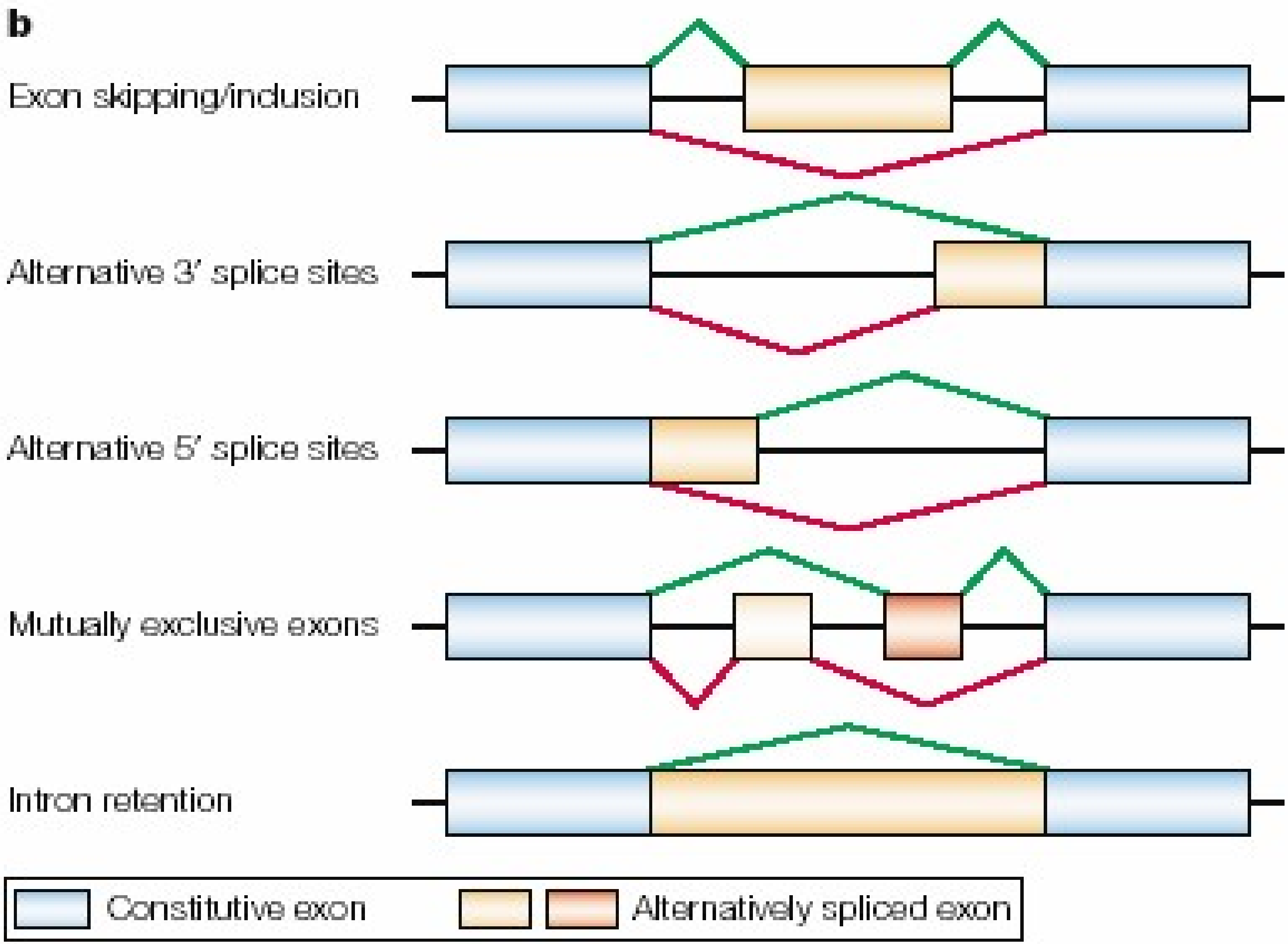


Figure 26.16

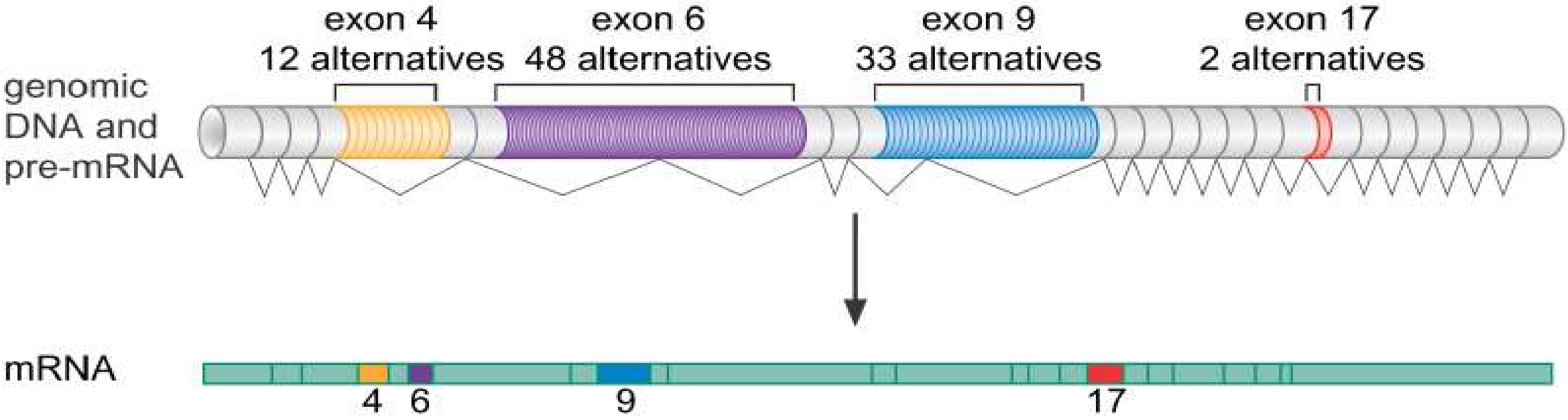
Alternative splicing is an important strategy to increase the number of gene products produced from a single gene.

Modes of alternative splicing.



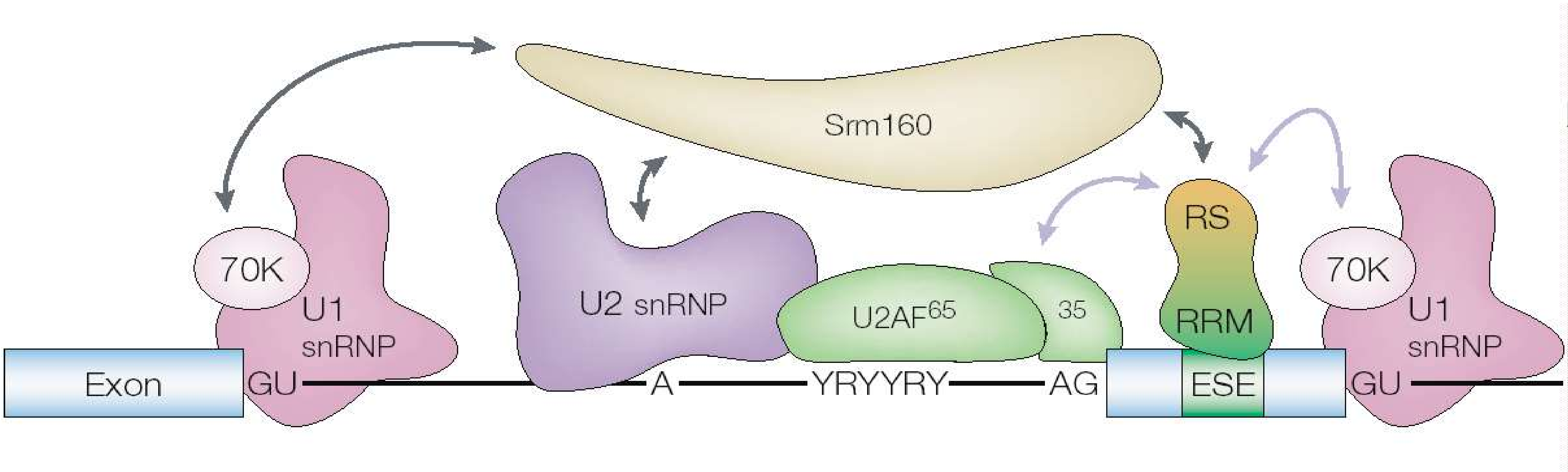
Cartegni L, Chew SL, Krainer AR.

Alternative splicing in *Drosophila* Dscam .



fg13-11

Inclusion of alternative exons may be promoted by exonic splicing enhancers that bind serine/arginine-rich (SR) RNA-binding proteins.



Heterogeneous ribonuclear proteins (hnRNP proteins) coat pre-mRNAs and mRNAs and often inhibit splicing events that SR proteins promote.
'Ying-yang' balance of these effects on many splicing events.

Splicing complexes Can be Formed by Intron Definition or Exon Definition

- The direct, **intron definition**, way of forming an E complex at short introns is for:
  - U1 snRNP to bind at the 5′ splice site
  - U2AF to bind at a pyrimidine tract between the branch site and the 3′ splice site

- Another possibility, **exon definition**, for very long introns is for the complex to form between:
  - U2AF at the pyrimidine tract
  - U1 snRNP at a downstream 5′ splice site

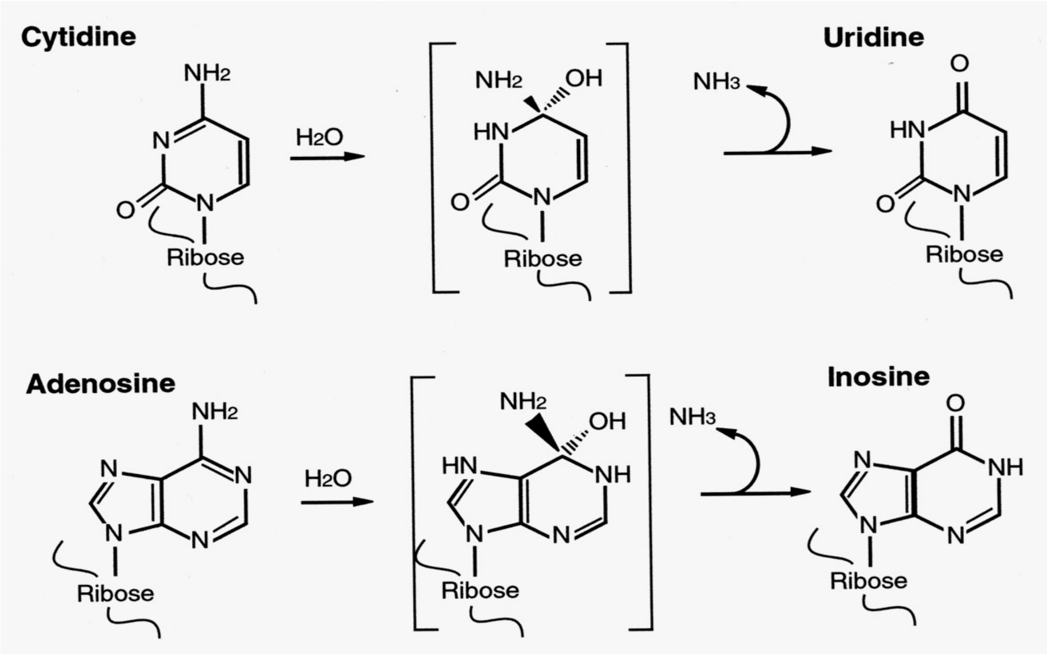
The importance of alternative splicing

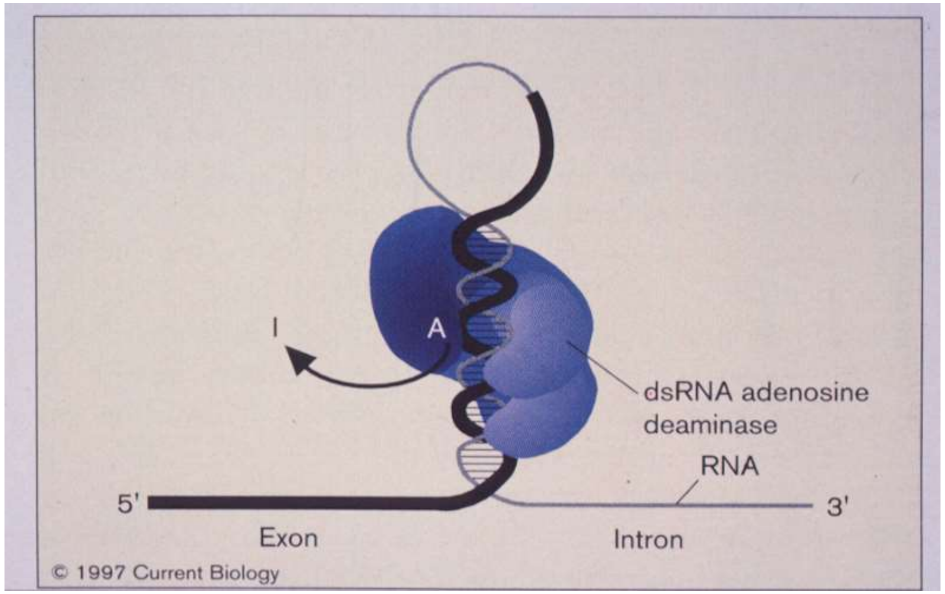
It is estimated that 95% of human genes are alternatively spliced from new sequencing information.

Alternative splicing creates protein diversity
26,000 human protein-encoding genes,
yet 90,000 proteins

Aberrant splicing is a frequent effect of human disease gene mutations.

Base deamination RNA editing also occurs site-specifically in some transcripts and changes codon meaning.





Adenosine deaminases acting on RNA (ADARs) edit 4% of *Drosophila* transcripts (972 sites), and ~250 known sites in human transcripts, including those encoding brain ion-channel subunits, particularly GluR B, the dominant subunit of AMPA class glutamate receptors.

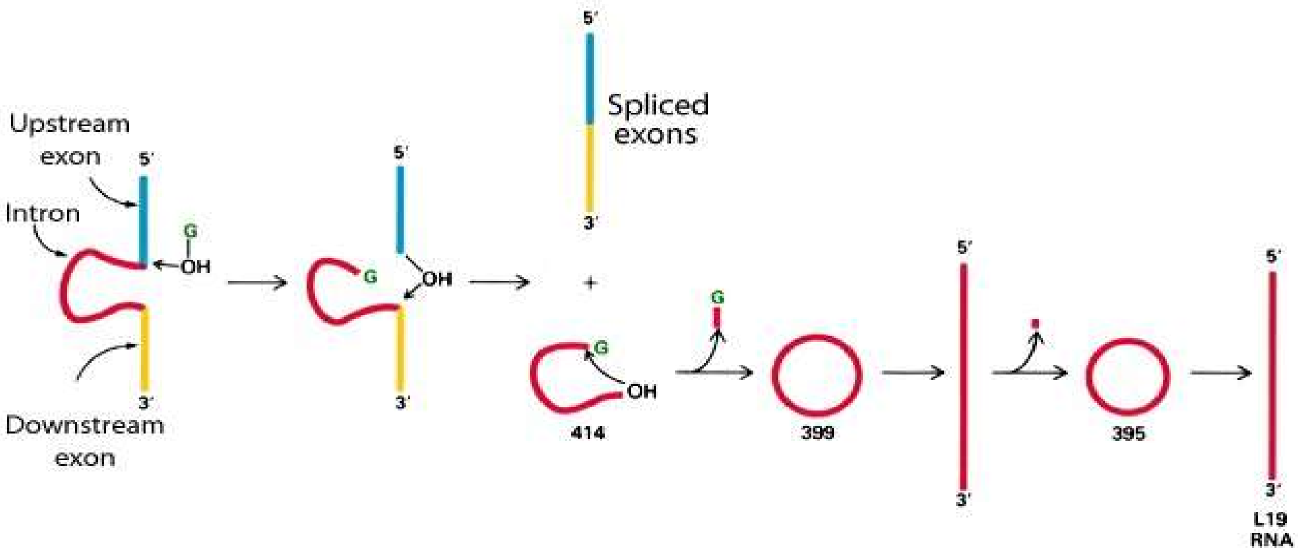
APOBECs are C to U deaminases first identified in RNA ediding but more important as DNA editors. Activation induced deaminase (AID) edits DNA in immunoglobulin genes to initiate Somatic Hypermutation of antibody variable regions. APOBEC3G edits HIV and interferes with virus replication.

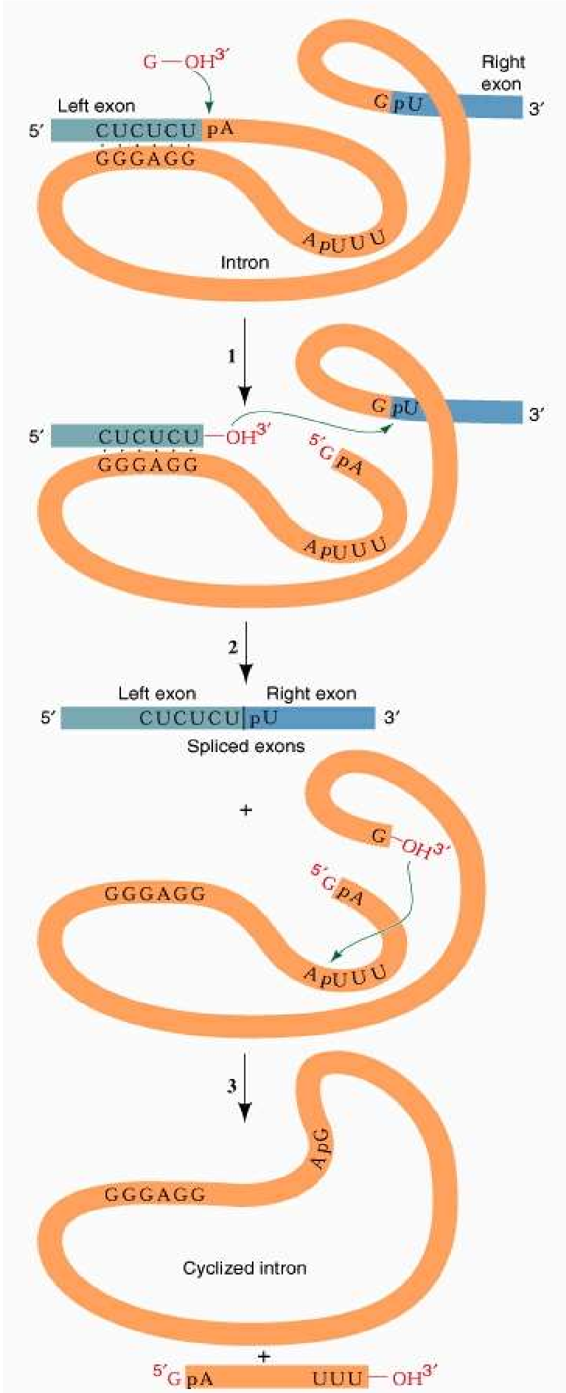
RNA catalysis, Ribozymes and the RNA World.

The RNA World.

Self-splicing introns are small transposable elements that remove themselves cleanly from the transcripts they are inserted in.

Group I intron splicing in 26S rRNA of *Tetrahymena* is catalysed by RNA in the absence of protein (Tom Cech).





The RNA World idea - RIBOZYMES.

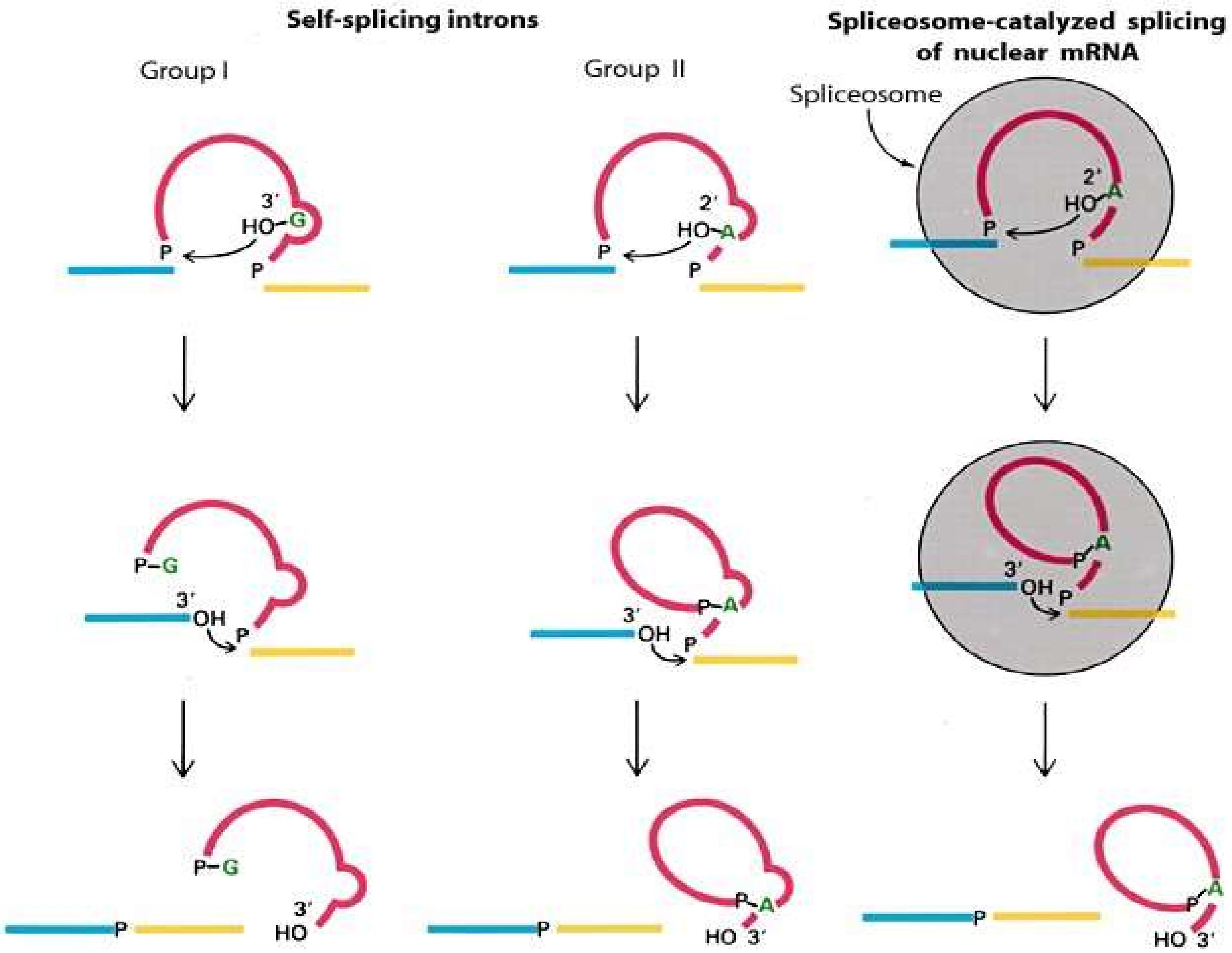
- Ribosomes and spliceosomes function through RNA-RNA interactions. Before the self-splicing introns were discovered it was assumed that catalysis would be caused by the proteins in these complexes.

- The discovery that RNA alone is catalytic in the Group I self-splicing intron showed that the RNA could be catalytic in ribosomes and spliceosomes also. RNA enzymes are called **RIBOZYMES** and many artificial ones have been created.

- This contributes to our understanding of possible origins of life as RNA alone could have carried out its own replication using natural ribozyme activities until translation was invented as a way to make more versatile protein catalysts.

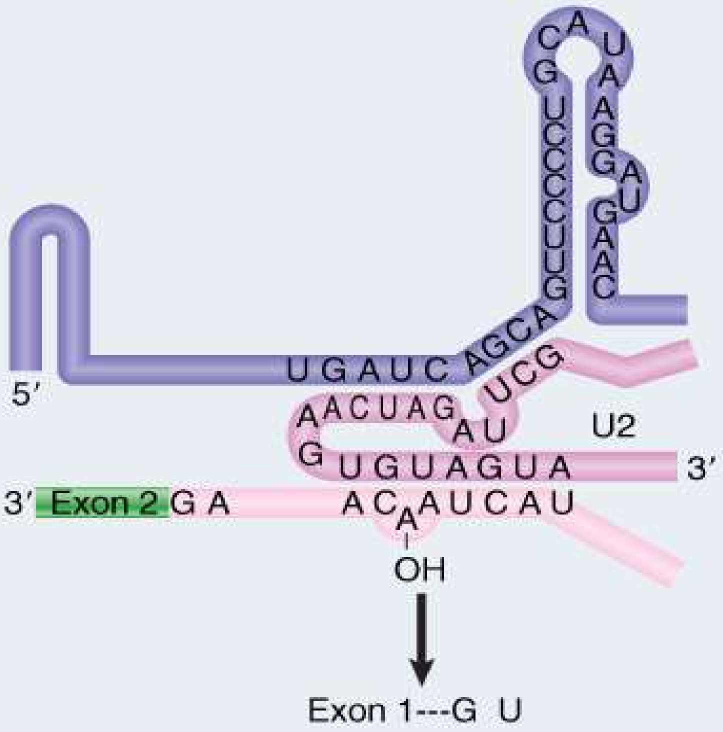
- DNA also may have been a later invention for use as a genomic library. It is more stable to hydrolytic cleavage than RNA and more readable by proteins than dsRNA to allow gene control.

The RNA World.
Group II introns also self-splice by RNA catalysis (Cech).



Nuclear and group II splicing are similar

Nuclear splicing constructs an active site from pairing between U6-U2 and U2-intron



Group II splicing constructs an active center from the base paired regions of domains 5 and 6

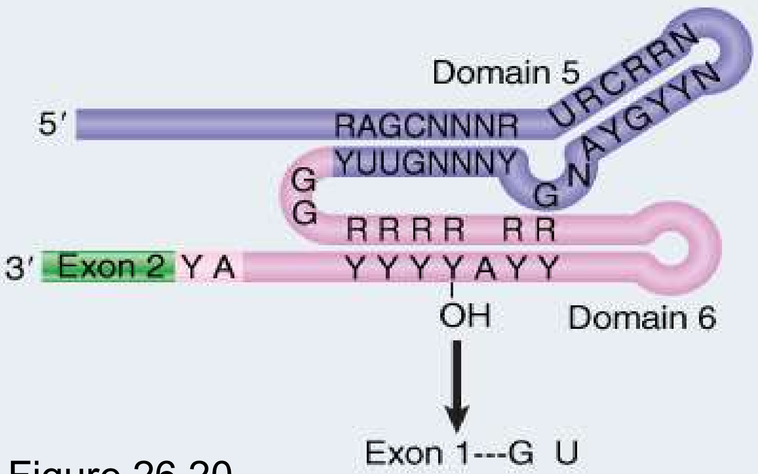


Figure 26.20

How introns might have spread.
Mobile Group II introns.

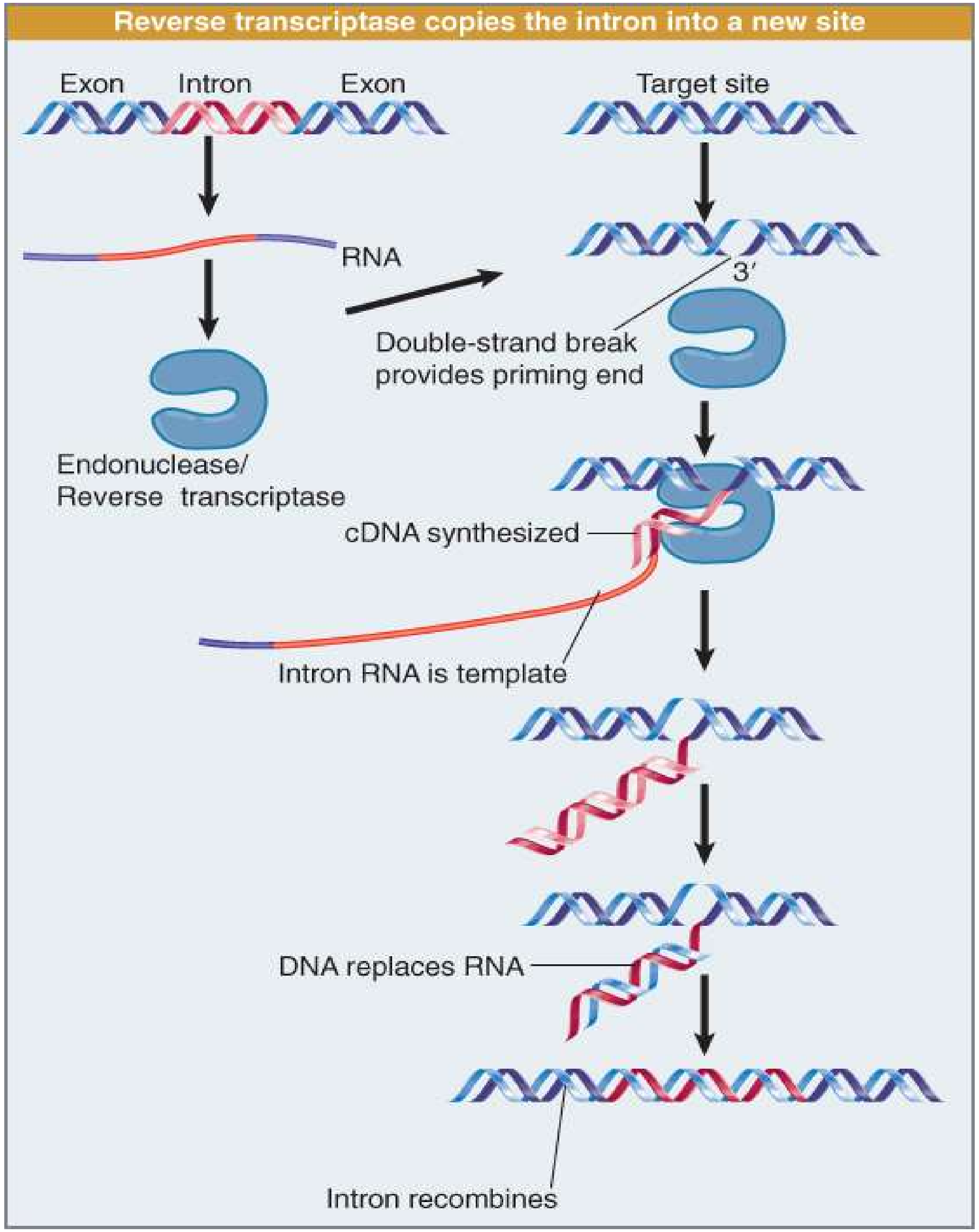
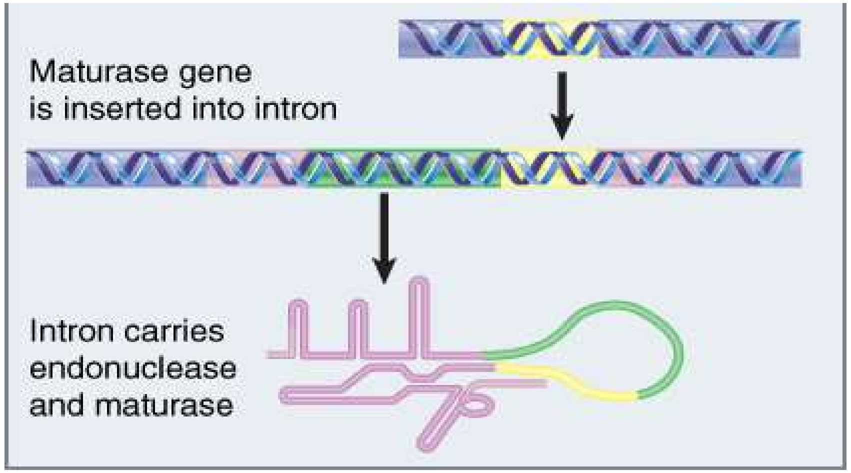


Figure 27.13

How spliceosomal introns might have evolved from self-splicing Group II introns.



- Autosplicing introns may require maturase protein activities encoded within the intron.
  - They assist folding into the active catalytic structure.
- Imagine that such an intron loses its maturase but can use maturase provided by other similar introns.
- This intron might eventually evolve to consist of just the target splice junctions GU..AG with all the catalytic activities supplied by a trans-acting splicesome.

Figure 27.15

Summary.

- Eukaryotic mRNA stability, 5’ cap and polyA addition and splicing.
- Mechanism of pre-mRNA splicing and alternative splicing. RNA editing.
- Self-splicing Group II introns, RNA catalysis, Ribozymes and the The RNA World.

Spares

Genome-wide locations of the coactivator P300/cEBP identify enhancers.

From: Nature. Author manuscript; available in PMC 2010 February 12.
Published in final edited form as:
Nature. 2009 February 12; 457(7231): 854–858.
doi: 10.1038/nature07730

Click on image to enlarge

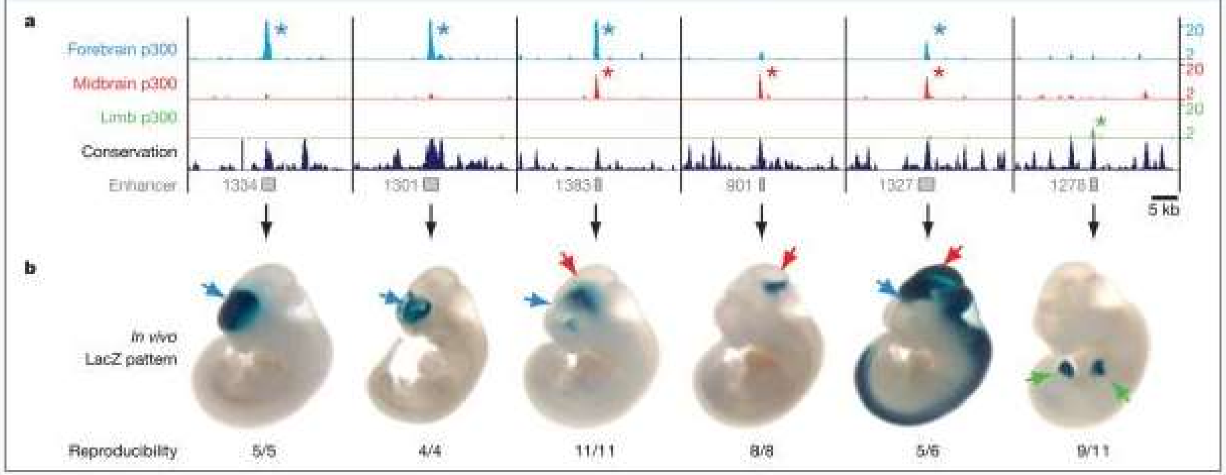


Figure 3

Examples of successful prediction of *in vivo* enhancers by p300 binding in embryonic tissues

a, Coverage by extended p300 reads in forebrain (blue), midbrain (red) and limb (green). Asterisks indicate significant (FDR < 0.01) p300 enrichment in chromatin isolated from the respective tissue. Multispecies vertebrate conservation plots (black) were obtained from the UCSC genome browser[50]. Grey boxes correspond to candidate enhancer regions. Numbers at the right indicate overlapping extended reads. b, Representative LacZ-stained embryos with *in vivo* enhancer activity at E11.5. Reproducible staining in forebrain, midbrain and limb is indicated by arrows. Numbers show the reproducibility of LacZ reporter staining. Additional embryos obtained with each construct and genomic coordinates are available using the enhancer ID in the bottom portion of **a** at the Vista Enhancer Browser[32].

Disease mutations in human enhancers.

- Until now human disease mutations in noncoding sequences were very difficult to identify.

- Once enhancers are identified across the human genome they can be included with exons and promoters in focused searches for disease-causing mutations.