13.19 Yeast Can Switch Silent and Active Mating-Type Loci

- The yeast mating-type locus $MAT$, a mating type cassette, has either the $MAT\alpha$ or $MAT\alpha$ genotype.
- Yeast with the dominant allele $HO$ switch their mating type at a high frequency.
- The allele at $MAT$ is called the active cassette.
- There are also two silent cassettes, $HML\alpha$ and $HMR\alpha$.

Figure 13.31: Changes of mating type occur when silent cassettes replace active cassettes of the opposite genotype.
13.19 Yeast Can Switch Silent and Active Mating-Type Loci

- Switching occurs if \( \text{MAT} \alpha \) is replaced by \( \text{HMR} \alpha \) or \( \text{MAT} \alpha \) is replaced by \( \text{HMR} \alpha \).

Figure 13.32: Silent cassettes have the same sequences as the corresponding active cassettes.
13.20 Unidirectional Gene Conversion Is Initiated by the Recipient $MAT$ Locus

- Mating-type switching is initiated by a double-strand break made at the $MAT$ locus by the HO endonuclease.
- The recombination event is a synthesis-dependent strand-annealing reaction.

Figure 13.34: Cassette substitution is initiated by a double-strand break in the recipient ($MAT$) locus.
Chip: Chromatin immunoprecipitation

• Fix the chromatin with formaldehyde to cross-link protein-DNA interactions
• Isolate chromatin and shear to 200-1000 bp
• Ab ppt POI-DNA cross links
• Wash and remove protein and DNA cross links
• Use pcr to amplify specific DNA sequences
Remodeling in Yeast \textit{HO} Gene Activation

• Chromatin immunoprecipitation (ChIP) can reveal the order of binding of factors to a gene during activation
• The \textit{HO} gene is activated just after mitosis resulting in mating-type switch
• As \textit{HO} gene is activated:
  – First factor to bind is Swi5
  – Followed by SWI/SNF and SAGA containing HAT Gcn5p
  – Next general transcription factors and other proteins bind
• Chromatin remodeling is among the first steps in activation of this gene
• Order could be different in other genes
Induction of IFN-B by RNA

ISG: interferon stimulated genes
8.4 Nucleosomes Are Covalently Modified

- Histones are modified by methylation, acetylation, phosphorylation, ubiquitylation, sumoylation, ADP-ribosylation, and other modifications.
- Combinations of specific histone modifications define the function of local regions of chromatin; this is known as the **histone code** hypothesis.

Figure 8.16: The histone tails can be acetylated, methylated, phosphorylated, and ubiquitylated at numerous sites.
Remodeling in the Human IFN-β Gene: The Histone Code

The Histone Code:

– The combination of histone modifications on a given nucleosome near a gene’s control region affects efficiency of that gene’s transcription
– This code is epigenetic, not affecting the base sequence of DNA itself

• Activators in the IFN-β enhanceosome can recruit a HAT (GCN5)
  – HAT acetylates some Lys on H3 and H4 in a nucleosome at the promoter
  – Protein kinase phosphorylates Ser on H3
  – This permits acetylation of another Lys on H3
Remodeling in the Human IFN-β Gene: TF Binding

- Remodeling allows TFIID to bind 2 acetylated lysines in the nucleosome through the dual bromodomain in TAF1
- TFIID binding
  - Bends the DNA
  - Moves remodeled nucleosome aside
  - Paves the way for transcription to begin
Histone code and activation of the IFN-B promoter

Fig. 13.29  Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.

Nucleosome free enhancer

Timing of histone acetylation in chromatin at the IFN-B promoter after virus infection

Fig. 13.27a

Post viral infection time points:

0 h 1 h 2 h 3 h 4 h 5 h 6 h 8 h 10 h 12 h 19 h 24 h

α-acH4 (K5, K8, K12, K16)

α-acH4 K8

α-acH4 K12

α-acH4 K5

α-acH4 K16

α-phH3 S10

α-acH3 (K9, K14)

α-acH3 (K9)

α-acH3 (K14)

α-TBP

INPUT

IFN-β m-RNA

© Reprinted from Cell v. 111, Agalioti et al., p. 383. © 2002, with permission from Elsevier Science
Heterochromatin

• **Euchromatin**: relatively extended and open chromatin that is potentially active

• **Heterochromatin**: very condensed with its DNA inaccessible
  – Microscopically appears as clumps in higher eukaryotes
  – Repressive character able to silence genes as much as 3 kb away
Heterochromatin is not active in transcription
Heterochromatin and Silencing

• Formation at the tips of yeast chromosomes (telomeres) with silencing of the genes is the telomere position effect (TPE)

• Depends on binding of proteins
  – RAP1 to telomeric DNA
  – Recruitment of proteins in this order:
    • SIR3
    • SIR4
    • SIR2
Silencing Information Regulator

SIR Proteins

• Heterochromatin at other locations in chromosome also depends on the SIR proteins

• SIR3 and SIR4 interact directly with histones H3 and H4 in nucleosomes
  – Acetylation of Lys 16 on H4 in nucleosomes prevents interaction with SIR3
  – Blocks heterochromatin formation

• Histone acetylation also works in this way to promote gene activity
Histone Methylation

• Methylation of Lys 9 in N-terminal tail of H3 attracts HP1
• This recruits a histone methyltransferase
  – Methylates Lys 9 on a neighboring nucleosome
  – Propagates the repressed, heterochromatic state
• Methylation of Lys and Arg side chains in core histones can have either repressive or activating effects
Model for involvement of histone methylation in chromatin repression

Fig. 13.32

H3K9-Me

Active, acetylated chromatin

Silenced heterochromatin spreads with HP1 addition and methylation

Histone Methylation

• Methylation of Lys 4 in N-terminal tail of H3 is generally tri-methylated (H3K4Me3) and is usually associated with the 5’-end of an active gene

• This modification appears to be a sign of transcription initiation

• Genome-wide ChIP analysis suggests that this may also play a role in controlling gene expression by controlling the re-starting of paused RNA polymerases
Summary

• Histone modifications can affect gene activity by two mechanisms:
• 1. By altering the way histone tails interact with DNA and with histone tails in neighboring nucleosomes, and thereby altering nucleosome cross-linking
• 2. By attracting proteins that can affect chromatin structure and activity
Ubiquitylation of histone H2B K123 is needed for methylation of H3 at K79 and K4

Rad6 is a ubiquitin ligase

Only K and not R can be ubiquitylated

\( \alpha \)-H3 K79Me

\( \alpha \)-H3 K4Me

\( \alpha \)-H3 K36Me

\( \alpha \)-H3
Modification Combinations

• Methylation occurs in a given nucleosome in combination with other histone modifications:
  – Acetylations
  – Phosphorylations
  – Ubiquitylations

• Each particular combination can send a different message to the cell about activation or repression of transcription

• One histone modification can also influence other, nearby modifications
Model for crosstalk among modifications on histone tails

Fig. 13.35
Nucleosomes and Transcription Elongation

• An important transcription elongation facilitator is FACT (facilitates chromatin transcription)
  – Composed of 2 subunits:
    • Spt16
      – Binds to H2A-H2B dimers
      – Has acid-rich C-terminus essential for these nucleosome remodeling activities
    • SSRP1 binds to H3-H4 tetramers
Nucleosomes and Transcription Elongation

• FACT facilitates transcription through a nucleosome by promoting loss of at least one H2A-H2B dimer from the nucleosome
• Also acts as a histone chaperone promoting re-addition of H2A-H2B dimer to a nucleosome that has lost such a dimer
mRNA Processing Events

• Most eukaryotic genes, in contrast to typical bacterial genes, are interrupted by noncoding DNA.
• RNA polymerases cannot distinguish the noncoding regions from the coding regions, so it transcribes everything.
• The cell must remove the noncoding RNA from the primary transcript via splicing.
• Eukaryotes also add special structures to the 5’ and 3’ ends of the transcript, called the cap and poly-A tail, respectively.
• All events occur in the nucleus before the mRNA emigrates to the cytoplasm.
14.1 Genes in Pieces

• Consider the sequence of the human $\beta$-globin gene as a sentence:
  
  This is *bhgty* the human $\beta$-globin *qwtzptlrbn* gene.

• Two italicized regions make no sense
  – Contain sequences unrelated to the globin coding sequences surrounding them
  – Intervening sequences, IVSs, or introns

• Parts of the gene making sense
  – Coding regions or exons

• Some lower eukaryotic genes have no introns
Evidence for Split Genes

• Most higher eukaryotic genes coding for mRNA, tRNA and a few coding for rRNA are interrupted by unrelated regions called introns

• Other parts of the gene, surrounding the introns, are called exons

• Exons contain the sequences that finally appear in the mature RNA product
  – Genes for mRNAs have been found with anywhere from 0 to 362 introns
  – tRNA genes have either 0 or 1 intron
RNA Splicing

• Introns are present in genes but not in mature RNA
• How does the information not find its way into mature RNA products of the genes?
  – Possibility 1: Introns are never transcribed
    • Polymerase somehow jumps from one exon to another
  – Possibility 2: Introns are transcribed
    • Primary transcript result, an overlarge gene product is cut down by removing introns
    • This is correct process
• The process of cutting introns out of immature RNAs and stitching together the exons to form the final product is RNA splicing
Splicing Outline

- Introns are transcribed along with exons in the primary transcript.
- Introns are removed as the exons are spliced together.

![Diagram showing the process of splicing]
Em visualization of precursor-spliced mRNA in an ‘R loop’

Figure 14.3

Globin mRNA precursor

Spliced globin mRNA

Tilghman, S., Curtis, P.; Tiemeier, P., and Weissmann, C. PNAS 75:1312, 1978
Simplified Mechanism of Splicing

- 2’-OH group of A within intron attacks the phosphodiester bond linking the first exon to the intron
- A lariat is formed due to the GU at the 5’ end of the intron forming a phosphodiester bond with the branchpoint A
- The free 3’ OH on exon 1 attacks the phosphodiester bond between the intron and exon 2
- The exons are then linked
Figure 14.6

Figure 14.10

Splicing intermediate

Hybridize to intron 5' end probe

Digest with RNase T1

Denature hybrid

+ other oligonucleotides

(oligo 46')

Digest with RNase T1
Lariat Branch Site

Lariat Branch site

Yeast: 5'UACUAAC3'

Eukaryote: 5'CACUGAC3'

2'-5' Linkage
Splice site sequence requirements

Splice site sequence requirements

\[
\begin{align*}
5' \text{exon} & \quad AG_{1} \quad GT \quad AG_{2} \quad \text{TT}XAG \quad ATC_{2}U \\
\text{intron} & \\
3' \text{exon} & \quad AC_{1}U
\end{align*}
\]

Conserved

Donor \quad AG_{1} \quad A_{G} \quad Acceptor
Spliceosomes

• Splicing takes place on a particle called a spliceosome

• Yeast spliceosomes and mammalian spliceosomes have sedimentation coefficients of 40S and 60S

• Spliceosomes contain the pre-mRNA
  – Along with snRNPs and protein splicing factors
  – These recognize key splicing signals and orchestrate the splicing process
Figure 14.12

Yeast Spliceosome

Wild-type pre-mRNA

Mutant pre-mRNA

Percent of total $^{32}$P

Fraction
Figure 14.13

Human spliceosome RNA content

Reprinted by permission of Elsevier Science
Alternative Splicing of Adenovirus E1A gene product

Figure 14.14

probe

611 nt
473 nt
136 nt

13S
12S
9S

RNase-protected products
Snurps or SnRNPs

• Small nuclear ribonuclear protein particles
• U1, U2, U4, U5 and U6 involved in nuclear RNA splicing
• Snurps constitute components of the spliceosome and are key components of splicing
• Each snurp has a unique role in splicing and the sequence of the RNA component of the snurp determines the role played in splicing
19.6 snRNAs Are Required for Splicing

- All the snRNPs except U6 contain a conserved sequence that binds the Sm proteins that are recognized by antibodies (anti-SM) generated in autoimmune disease.

- **splicing factor** – A protein component of the spliceosome that is not part of one of the snRNPs.

- **transesterification** – A reaction that breaks and makes chemical bonds in a coordinated transfer so that no energy is required.
19.7 Commitment of Pre-mRNA to the Splicing Pathway

• U1 snRNP initiates splicing by binding to the 5’ splice site by means of an RNA–RNA pairing reaction.

• The commitment complex (or E complex) contains U1 snRNP bound at the 5’ splice site and the protein U2AF bound to a pyrimidine tract between the branch site and the 3’ splice site.
Figure 19.0.8.0: U1 snRNA has a base-paired structure that creates several domains. The 5’ end remains single stranded and can base pair with the 5’ splice site.
Yeast splicing of mRNA

Most intron-containing genes are interrupted by a single small 100-300 nucleotide intron

Splicing does not involve the SR proteins

U1, BBP and Mud2 bind simultaneously to the 5’ and 3’ splice sites
Figure 14.36

Yeast

Mammals
3’-Splice Site Selection

• Splicing factor Slu7 is required for correct 3’-splicing site selection
• Without Slu7, splicing to correct 3’-splice site AG is suppressed and splicing to aberrant AG’s within 30 nt of the branchpoint is activated
• U2AF is also required for 3’-splice site recognition
• 65-kD U2AF subunit binds to polypyrimidine tract upstream of 3’-splice site and 35-kD subunit binds to the 3’-splice site AG
19.7 Commitment of Pre-mRNA to the Splicing Pathway

• In cells of multicellular eukaryotes, SR proteins play an essential role in initiating the formation of the commitment complex.

• Pairing splice sites can be accomplished by intron definition or exon definition.
19.7 Commitment of Pre-mRNA to the Splicing Pathway

Figure 19.10: The commitment (E) complex formation.
19.7 Commitment of Pre-mRNA to the Splicing Pathway

Figure 19.11: The two routes for initial recognition of 5’ and 3’ splice sites are intron definition and exon definition.
19.8 The Spliceosome Assembly Pathway

- The commitment complex progresses to prespliceosome (the **A complex**) in the presence of ATP.
- Recruitment of U5 and U4/U6 snRNPs converts the A complex to the mature spliceosome (the B1 complex).
- The B1 complex is next converted to the B2 complex, in which U1 snRNP is released to allow U6 snRNA to interact with the 5′ splice site.
19.8 The Spliceosome Assembly Pathway

- When U4 dissociates from U6 snRNP, U6 snRNA can pair with U2 snRNA to form the catalytic active site.
- Both transesterification reactions take place in the activated spliceosome (the C complex).
- The splicing reaction is reversible at all steps.

Figure 19.12: The splicing reaction proceeds through discrete stages.
Figure 19.0.13.0: U6/U4 pairing is incompatible with U6/U2 pairing. When U6 joins the spliceosome it is paired with U4. Release of U4 allows a conformational change in U6; one part of the released sequence forms a hairpin and the other part pairs with U2. An adjacent region of U2 is already paired with the branch site, which brings U6 into juxtaposition with the branch. Note that the substrate RNA is reversed from the usual orientation and is shown 3’ to 5’.
19.10 Pre-mRNA Splicing Likely Shares the Mechanism with Group II Autocatalytic Introns

Figure 19.15: Three classes of splicing reactions proceed by two transesterifications.
Figure 19.0.16.0: Nuclear splicing and group II splicing involve the formation of similar secondary structures. The sequences are more specific in nuclear splicing; group II splicing uses positions that may be occupied by either purine (R) or pyrimidine (Y).
U2 U6 and U5 RNA interactions

Figure 14.26
19.11 Splicing Is Temporally and Functionally Coupled with Multiple Steps in Gene Expression

- Splicing can occur during or after transcription.
- The transcription and splicing machineries are physically and functionally integrated.
- Splicing is connected to mRNA export and stability control.
- **exon junction complex (EJC)** – A protein complex that assembles at exon–exon junctions during splicing and assists in RNA transport, localization, and degradation.
Figure 19.18: The EJC (exon junction complex) is deposited near the splice junction as a consequence of the splicing reaction.
Role of the RNA Polymerase II CTD

- C-terminal domain of the Rpb1 subunit of RNA polymerase II stimulates splicing of substrates that use exon definition
- This does not apply to those that use intron definition to prepare for splicing
- CTD binds to splicing factors and could assemble the factors at the end of exons to set them off for splicing
Fig. 14.34  Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.

(a) DNA

m^7G

Splicing factors

CTD

(b)

m^7G

Intron

(c)

m^7G

Alternative Splicing

• Transcripts of many eukaryotic genes are subject to alternative splicing
  – This splicing can have profound effects on the protein products of a gene
  – Can make a difference between:
    • Secreted or membrane-bound protein
    • Activity and inactivity

• Products of 3 genes in sex determination pathway of the fruit fly are subject to alternative splicing
Alternative Splicing: Immunoglobulin genes

Figure 14.37

Secreted $\mu$ mRNA

Membrane $\mu$ mRNA
Tra and Tra-2

• Tra and its partner Tra-2 act in conjunction with one or more other SR proteins to commit splicing at the female-specific splice site on the dsx pre-mRNA

• Commitment is probably the basis of most, if not all, alternative splicing schemes
Alternative Splicing Patterns

• Alternative splicing of the same pre-mRNA gives rise to very different products
  – Alternative splicing patterns occur in over half of human genes
  – Many genes have more than 2 splicing patterns, some have thousands
Control of Splicing

- Begin transcripts at alternative promoters
- Some exons can simply be ignored resulting in deletion of the exon
- Alternative 5’-splice sites can lead to inclusion or deletion of part of an exon
- Alternative 3’-splice sites can lead to inclusion or deletion of part of an exon
- A retained intron can be retained in the mRNA if it is not recognized as an intron
- Polyadenylation causes cleavage of pre-mRNA and loss of downstream exons
Silencing of Splicing

- What stimulates recognition of signals under only some circumstances?
- Exons can contain sequences –
  - Exonic splicing enhancers (ESEs) stimulate splicing
  - Exonic splicing silencers (ESSs) inhibit splicing
Alternative Splicing Summary

• Alternative splicing is very common in higher eukaryotes

• It represents a way to get more than one protein product out of the same gene and a way to control gene expression in cells

• Such control is exerted by splicing factors that bind to splice sites and a branchpoint, and also by proteins that interact with ESEs, ESSs and intronic splicing elements
14.3 Self-Splicing RNAs

• Some RNAs could splice themselves without aid from a spliceosome or any other protein
• *Tetrahymena* 26S rRNA gene has an intron, splices itself in vitro
  – Group I introns are a group of self-splicing RNAs
  – Another group, Group II introns also have some self-splicing members
Group I Introns self splice

- Tetrahymena 26S rRNA precursor can be removed in vitro with no protein addition to the RNA.
- The reaction requires a guanine nucleotide on the 5’ splice site, adding G to the 5’ end of the intron and releasing the first exon.
- The first exon attacks the 3’ splice site ligating the two exons together and releasing the linear intron.
Group I Introns

• Group I introns can be removed in vitro with no help from protein

• Reaction begins with attack by a guanine nucleotide on the 5’-splice site
  – Adds G to the 5’-end of the intron
  – Releases the first exon

• Second step, first exon attacks the 3’-splice site
  – Ligates 2 exons together
  – Releases the linear intron

• Intron cyclizes twice, losing nucleotides each time, then linearizes a last time
Group I Intron: *Tetrahymena* 26S rRNA precursor

Fig. 14.43

Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.

(a) Pre-rRNA

(b) Exon 1

Intron

Exon 2

GpA

GOH

Intron

Exon 1

UOH

GpA

Exon 2

GpU

Intron

Exon 1

UOH

GpU

Exon 2

GpA

GOH
GMP held in the pocket of the *Tetrahymena* 26S rRNA

---

What happens to the intron?
Mechanism of tRNA splicing

- The intron is removed by a membrane bound endonuclease leaving a 2’-3’ phosphate group
- The 5’ end of the 3’ half of the tRNA is phosphorylated by a kinase and the end is further modified by ligase to add an AMP molecule
- The two half-molecules are ligated through a 3’-5’ phosphodiester linkage
- Finally the 2’ phosphate group is removed and a mature tRNA is formed
Figure 14.51

Mechanism of tRNA splicing

- **Endonuclease (a)**
  - **Pre-tRNA**
  - **2' O**
  - **5'**
  - **3' O**
  - **HO**
  - **5'**
  - **3'**

- **Kinase + ATP (c)**
  - **P**
  - **5'**
  - **3'**

- **Ligase + ATP (d)**
  - **PP_i**
  - **P**
  - **A**

- **Ligase (e)**
  - **P**
  - **A**

- **Phosphatase (g)**
  - **P**
  - **OH**

- **Mature tRNA**
  - **5'**
  - **3'**
Precursor mRNAs are “capped”

- Soon after polymerase clearance of the promoter a CTD bound enzyme guanylylmethyltransferase adds an inverted GMP molecule to the 5’ end of the transcript
Figure 15.3
Cap Synthesis

- First step
  - RNA triphosphatase removes terminal phosphate from pre-mRNA
  - Then, guanylyl transferase adds capping GMP from GTP

- Next, 2 methyl transferases methylate $N^7$ of capping guanosine and 2'-O-methyl group of penultimate nucleotide

- This occurs early in transcription, before chain is 30 nt long
Figure 15.4
Functions of Caps

- Protect mRNA from degradation
- Enhance translation of the mRNA
- Enhance the nuclear export of the mRNA
- Enhance the efficiency of splicing
Poly AMP addition to the 3’ end of the mRNA

• Poly (A) polymerase adds approximately 250 AMP residues to the transcript
• Poly (A) increases the lifetime of the mRNA and its ability to be translated
• A specific recognition sequence AAUAAA occurs approximately 20 bases before the poly (A) site
Basic Mechanism of Polyadenylation

- Transcription of eukaryotic genes extends beyond the polyadenylation site
- The transcript is:
  - Cleaved
  - Polyadenylated at 3’-end created by cleavage
Polyadenylation Signals

• An efficient mammalian polyadenylation signal consists of:
  – AAUAAA motif about 20 nt upstream of a polyadenylation site in a pre-mRNA
  – Followed 23 or 24 bp later by GU-rich motif
  – Followed immediately by a U-rich motif

• Variations on this theme occur in nature
  – Results in variation in efficiency of polyadenylation
  – Plant polyadenylation signals usually contain AAUAAA motif
  – More variation exists in plant than in animal motif
  – Yeast polyadenylation signals are even more different
Cleavage of Pre-mRNA

• Polyadenylation involves both:
  – Pre-mRNA cleavage
  – Polyadenylation at the cleavage site

• Cleavage in mammals requires several proteins
  – CPSF – cleavage and polyadenylation specificity factor
  – CstF – cleavage stimulation factor
  – CF I
  – CF II
  – Poly (A) polymerase
  – RNA polymerase II
Initiation of Polyadenylation

- Short RNAs mimic a newly created mRNA 3’-end can be polyadenylated
- Optimal signal for initiation of such polyadenylation of a cleaved substrate is AAUAAA followed by at least 8 nt
- When poly(A) reaches about 10 nt in length, further polyadenylation becomes independent of AAUAAA signal and depends on the poly(A) itself
- 2 proteins participate in the initiation process
  - Poly(A) polymerase
  - CPSF binds to the AAUAAA motif
Elongation of Poly(A)

• Elongation of poly(A) in mammals requires a specificity factor called poly(A)-binding protein II (PAB II)

• This protein
  – Binds to a preinitiated oligo(A)
  – Aids poly(A) polymerase in elongating poly(A) to 250 nt or more

• PAB II acts independently of AAUAAA motif
  – Depends only on poly(A)
  – Activity enhanced by CPSF
Model for Polyadenylation

- Factors assemble on the pre-mRNA guided by motifs
- Cleavage occurs
- Polymerase initiates poly(A) synthesis
- PAB II allows rapid extension of the oligo(A) to full-length
Figure 15.26

Cleavage and poly(A) addition

5'-GpppG

AAUAAA

PAP

82

100

30

73

160

CPSF

50

77

64

G/U

CstF

RNA Pol II

3'

CF I/CF II
Poly(A) Polymerase

- Cloning and sequencing cDNAs encoding calf thymus poly(A) polymerase reveal a mixture of 5 cDNAs derived from alternative splicing and alternative polyadenylation.
- Structures of the enzymes predicted from the longest sequence includes:
  - RNA-binding domain
  - Polymerase module
  - 2 nuclear localization signals
  - Ser/Thr-rich region – this is dispensable for activity in vitro
Turnover of Poly(A)

- Poly(A) turns over in the cytoplasm
- RNases tear it down
- Poly(A) polymerase builds it back up
- When poly(A) is gone mRNA is slated for destruction
15.3 Coordination of mRNA Processing Events

- After reviewing capping, polyadenylation and splicing, it is clear that these processes are related
- Cap can be essential for splicing, but only for splicing the first intron
- Poly(A) can also be essential, but only for splicing out the last intron
Processing occurs during Transcription

All three of the mRNA-processing events take place during transcription

– Splicing begins when transcription is still underway

– Capping
  • When nascent mRNA is about 30 nt long
  • When 5’-end of RNA first emerges from polymerase

– Polyadenylation occurs when the still-growing mRNA is cut at the polyadenylation site
Binding of CTD of Rpb1 to mRNA-Processing Proteins

• The CTD of Rpb1 subunit of RNA polymerase II is involved in all three types of processing

• Capping, polyadenylating, and splicing enzymes bind directly to the CTD which serves as a platform for all three activities
CTD Phosphorylation

• Phosphorylation state of the CTD of Rpb1 in transcription complexes in yeast changes as transcription progresses
  – Transcription complexes close to the promoter contain phosphorylated Ser-5
  – Complexes farther from the promoter contain phosphorylated Ser-2

• Spectrum of proteins associated with the CTD also changes
  – Capping guanylyl transferase is present early when the complex is close to promoter, not later
  – Polyadenylation factor Hrp1 is present in transcription complexes near and remote from promoter
RNA Processing Organized by CTD

Fig. 15.32

(c) Cleavage and polyadenylation complex

Further phosphorylation of CTD

Splicing complexes

Capping complex

Nascent RNA

m^7G

Mechanism of Termination

• Termination of transcription by RNA polymerase II occurs in 2 steps:
  – Transcript experiences a cotranscriptional cleavage (CoTC) within termination region downstream of the polyadenylation site
    • This occurs before cleavage and polyadenylation at the poly(A) site
    • It is independent of that process
  – Cleavage and polyadenylation occur at the poly(A) site
    • Signals polymerase to dissociate from template
Termination Signal

• CoTC element downstream of the polyadenylation site in the human β-globin mRNA is a ribozyme that cleaves itself
  – This generates a free RNA 5’-end
  – This cleavage is required for normal transcription termination
  – It provides an entry site for Xrn2, a 5’→3’ exonuclease that loads onto the RNA and chases RNA polymerase by degrading the RNA
Xrn2, Exonuclease

• Xrn2 terminates transcription like a “torpedo”
• There is a similar torpedo mechanism in yeast where cleavage at poly(A) site provides entry for the 5’ → 3’ exonuclease Rat1
• Rat1 degrades the RNA until it catches the polymerase and terminates transcription
Torpedo Model for Transcription Termination
Role of Polyadenylation in mRNA Transport

• Polyadenylation is required for efficient transport of mRNAs from their point of origin in the nucleus to the cytoplasm
Poly (A) consensus sequence

Consensus

$A_{98}A_{86}U_{98}A_{98}A_{95}A_{96}$

$U_{12}$

Polyadenylation activity

AAUAAAG
AAUUGAA
AAUAAGA
AAUAAAC
AAUCAAA
AAUCGAA
AAUGGAA
AAGAANA
AAUGAAA
AAUGUAA
AAUAGA
AAUCUAA
ACUAAA
GUAAN
GUAA
UAAN
CAAA

Figure 15.19
### Table 15.2 Mammalian Factors Required for 3′-cleavage and Polyadenylation

<table>
<thead>
<tr>
<th>Factor</th>
<th>Polypeptides (kD)</th>
<th>Properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poly(A) polymerase (PAP)</td>
<td>82</td>
<td>Required for cleavage and polyadenylation; catalyzes poly(A) synthesis</td>
</tr>
<tr>
<td>Cleavage and polyadenylation specificity factor (CPSF)</td>
<td>160</td>
<td>Required for cleavage and polyadenylation; binds AAUAAA and interacts with PAP and CstF</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td></td>
</tr>
<tr>
<td></td>
<td>70</td>
<td></td>
</tr>
<tr>
<td></td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>Cleavage stimulation factor (CstF)</td>
<td>77</td>
<td>Required only for cleavage; binds the downstream element and interacts with CPSF</td>
</tr>
<tr>
<td></td>
<td>64</td>
<td></td>
</tr>
<tr>
<td></td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>Cleavage factor I (CF I)</td>
<td>68</td>
<td>Required only for cleavage; binds RNA</td>
</tr>
<tr>
<td></td>
<td>59</td>
<td></td>
</tr>
<tr>
<td></td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>Cleavage factor II (CF II)</td>
<td>unknown</td>
<td>Required only for cleavage</td>
</tr>
<tr>
<td>RNA polymerase II (especially CTD)</td>
<td>many</td>
<td>Required only for cleavage</td>
</tr>
<tr>
<td>Poly(A)-binding protein II (PAB II)</td>
<td>33</td>
<td>Stimulates poly(A) elongation; binds growing poly(A) tail; essential for poly(A) tail length control</td>
</tr>
</tbody>
</table>