Processing of RNA II
Biochemistry 302

February 14, 2005
Bob Kelm
What’s an intron?

- Transcribed sequence removed during the process of mRNA maturation (non protein-coding sequence)
- Discovered by P. Sharp and R. Roberts in late 1970s by EM analysis of genomic DNA-mRNA (poly A) hybrids encoding the major virion capsid protein, hexon
  - DNA loops correspond to introns
  - Same looping pattern later shown w/ ovalbumin mRNA-genomic DNA hybrids

Chicken ovalbumin
Overview of intron excision or exon splicing mechanisms (all 2-step)

- **Group I Introns**
  - Cofactor guanosine (3’OH) attacks exon/5’ intron boundary.
  - “Linear” intervening sequence (IVS) is generated.

- **Group II Introns**
  - Branch point adenosine (2’OH) attacks exon/5’ intron boundary.
  - Lariat-like IVS is generated.

- **Nuclear mRNA splicing**
  - Branch point adenosine (2’OH) attacks exon/5’ intron boundary generating lariat IVS.
  - Requires assistance of snRNPS.

- **Nuclear tRNA “splicing”**
  - Protein (not RNA)-dependent.
  - Requires endonuclease & ligase.

*Figure 2* Splicing mechanisms of the four major groups of precursor RNAs. Wavy lines indicate introns. Smooth lines indicate flanking exons. For nuclear
### Consensus sequences near 5’ and 3’ splice sites in vertebrate pre-mRNAs

**Table 28-6**

<table>
<thead>
<tr>
<th>Protein, Intron</th>
<th>5’ E1 splice site</th>
<th>Intron</th>
<th>E2 splice site</th>
<th>3’</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ovalbumin, intron 3</td>
<td>⋯UCAG</td>
<td>GUACAG</td>
<td>⋯UGUAUUC</td>
<td>AG</td>
</tr>
<tr>
<td>β globin, human, intron 1</td>
<td>⋯CGAG</td>
<td>GUUGGU</td>
<td>⋯CACCCU</td>
<td>UAG</td>
</tr>
<tr>
<td>β globin, human, intron 2</td>
<td>⋯CAGG</td>
<td>GUGAGU</td>
<td>⋯CCUCCA</td>
<td>CAG</td>
</tr>
<tr>
<td>Immunoglobulin I, L-VI</td>
<td>⋯UCAG</td>
<td>GUCAGC</td>
<td>⋯UGUUCG</td>
<td>AG</td>
</tr>
<tr>
<td>Rat preproinsulin</td>
<td>⋯CAAG</td>
<td>GUAAGC</td>
<td>⋯CCCUGG</td>
<td>CAAG</td>
</tr>
<tr>
<td>Consensus sequences*</td>
<td>—AG</td>
<td>GURAGY</td>
<td>⋯YYYYY</td>
<td>—AG</td>
</tr>
</tbody>
</table>

*Here R stands for purine and Y for pyrimidine. Residues listed for the consensus sequence are those found in two-thirds or more of over 100 cases analyzed. The residues shown in red are invariant in all cases analyzed.*

---

![Diagram of pre-mRNA structure](image)

*A U

Lodish et al. Molecular Cell Biology, 3rd ed.

*A C

*Minor class: 1/300 introns but not in yeast or worms (J. Steitz)
Basic chemistry of exon splicing → phosphate ester bond exchange or transesterification

Lodish et al. Molecular Cell Biology, 3rd ed.
RNA splicing can be studied in vitro (from M. Green and coworkers)
Excision of introns from nuclear pre-mRNAs is spliceosome-mediated

- snRNAs (U-rich)
  - U1.....U6 (U3 nucleolus)
  - Small: 107 to 210 nt w/ lots of secondary structure
  - Some transcribed by RNAP II, others by RNAP III
- snRNPs (“snurps”)
  - Some common to all snRNAs (Sm binds U1, U2, U4, U5).
  - Some specific to individual snRNAs.
- Other specialty proteins
  - RNA helicases (DEAD box family)
  - SR proteins

Fig. 28-31

U1 interacts w/ 5’ intron-exon boundary

U2 interacts w/ 3’ intron-exon boundary. Note how “bulge” may serve to activate adenosine 2’OH.
Putative mechanism of intron removal from pre-mRNAs (orienting splice sites)

Splice site consensus:

5′ = G:GU

exon:intron

3′ = Py10U/CAG:C

intron:exon

Note: U1 and U2 snRNPs are necessary but not sufficient to mediate splicing.

Fig. 28-32
Model for snRNP-mediated splicing of nuclear pre-mRNA exons (U1/2-dependent)

1: U1 snRNP binds to 5′ splice site through base pairing with 5′ end of U1 RNA

2: Auxiliary factor U2AF binds poly Pyr tract to assist binding of U2 snRNP to branch point sequence (ATP-dependent). U1 and U2 facilitate loop structure.

3: Spliceosome assembles (60S): U4/U6 & U5 on U1 + U2 snRNPs

4,5: Internal rearrangement and probable U1/U4 dissociation then U2/U6 and U5-dependent lariat formation and exon splicing via 5′ and 3′ consensus sites. U2:U6 interaction is crucial. U6 is the putative ribozyme.

6-8: Spliceosome disassembly releasing ligated exons, intron lariat & U snRNPs (recycled).
Pre-mRNA splicing is probably a co-transcriptional process

Lehninger Principles of Biochemistry, 4th ed., Ch 26
Mechanisms to generate alternative transcripts from a single gene

Why....
- Economize genomic size
- To allow tissue/cell-specific gene expression

How....
- Alternative promoter usage: mouse amylase
- Poly A site selection: Ig heavy chain class switching
- Differential RNA splicing: muscle-specific genes (e.g. \(\alpha\)-tropomyosin, troponins, SM myosin)
- Alternative polyA + splicing: dictates calcitonin or CGRP hormone expression in thyroid versus brain
- RNA editing: apoB

Lehninger Principles of Biochemistry, 4th ed., Ch 26
Alternative splicing for cell-type specific expression of mRNAs

Mechanism: Cell type-specific proteins interact with spliceosome to alter the default choice of splice site selection.

**known**

- Striated muscle
- Striated muscle'
- Myoblast
- Smooth muscle

**Inferred (nuclease protection)**

- Nonmuscle/fibroblast
- Hepatoma
- Brain

**default product**

Fig. 28-34
Cis-elements and other RNA-binding proteins regulate alternative splicing

- **Cis-elements**
  - Exon splicing enhancers (ESEs) – purine-rich RNA sequences that stimulate splicing of adjacent 5’ introns
  - Intron enhancer sequences
  - Negative regulatory elements (found in both introns and exons)

- **Trans-acting proteins**
  - SR proteins – family of RNA-binding proteins required for spliceosome assembly and the activity of distinct ESEs, regulated by SR protein kinases
  - Sex and tissue-specific splicing factors (*Drosophila* sex determination)
Group I intron: self-splicing RNA

- Cech and coworkers (1982)
  - Discovered a 414 nt self-splicing rRNA intron
  - *Tetrahymena thermophila* (ciliated protozoan, unicellular animal model)

- Highly structured RNA
  - Self-splicing requires Mg$^{2+}$ (~2 mM) and a guanosine or guanylate cofactor.
  - RNA modified during reaction (so not a real catalyst).

- Two transesterification rxns
  - 1$^{st}$: guanosine nucleoside cofactor attacks 5$'$ splice site
  - 2$^{nd}$: 3$'$ end (OH) of 1st exon attacks G at the 3$'$ splice site

Further truncation of the released IVS yields a catalytic 395 nt RNA molecule known as L-19 IVS. It can shorten or elongate some oligonucleotides in vitro.
Group I intron from *Tetrahymena* is preorganized to bind substrates

P1 domain containing the 5’ splice site is thought to fit here.

Self-splicing: Group II introns

- Found in mitochondrial and chloroplast pre-rRNAs.
- Fold into conserved secondary structure characterized by numerous stem loops.....similar to spliceosomal U snRNAs.
- Self-cleavage mechanism similar to spliceosome but only happens under conditions of elevated temperature and [Mg$^{2+}$] \textit{in vitro}.
- U snRNAs may have evolved from group II introns.
- Some can behave as mobile elements in the genome.
- Encode maturases, proteins required for facilitating genetic transposition and/or splicing of group I and II introns \textit{in vivo}.

Lodish et al. Molecular Cell Biology, 3rd ed.
Lehninger Principles of Biochemistry, 4th ed., Ch 26