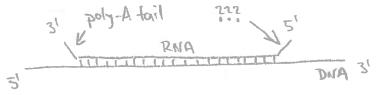
5. RNA Splicing I

5.1. Introduction to RNA Splicing

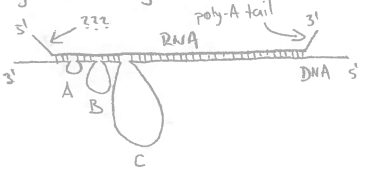
- process, where long pre-mRNAs are spliced to [remove introns / retain exons] ← cutting & joining of RNA strands
- processed mRNAs are used as templates for tnl
- function [make multiple variants of gene product / control gene expression]
- very common in eukaryotes, occurs in nucleus

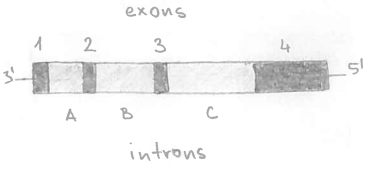
5.2. Discovery of Splicing and R-Loop Mapping

- from studies of txn of adenovirus (takes over euk. txn → viral RNAs very abundant)
- R loop mapping [purify RNA from viral infected cells / hybridize isolated RNA w/ DNA fragment from adenovirus genome / observe hybridization using electron microscopy]
  - small DNA fragment



  - longer DNA fragment



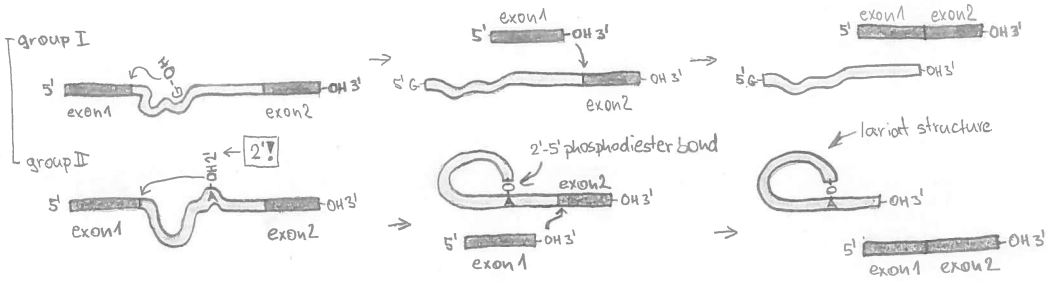


5.3. Types of Splicing Reactions

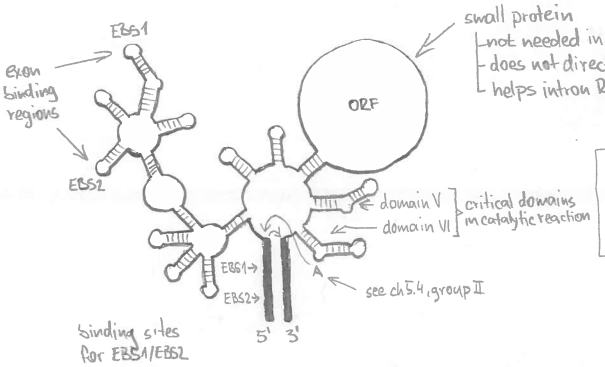
- self-splicing (many mechanistic details of how the RNA participates in the splicing reaction have been gained from studying self-splicing RNAs)
  - group I
  - group II
- spliceosome-mediated
  - U2 spliceosome (mediates splicing of most euk. protein-encoding mRNAs)
  - U12 spliceosome
- more specialized

5.4. Self-Splicing Introns, Mechanistic Overview

- not very common
- historically important: studying [RNA structure / role of RNA in promoting catalysis]
- under certain conditions, RNAs can do their splicing reaction on their own (only some divalent metal ions and salts are needed) ← Mg^{2+}, Mn^{2+}, ... ions
  - ↳ introns themselves are catalyzing the reaction ← ribozymes: "enzymes" made of RNA
  - all reactions occuring are transesterification reactions → isoenergetic ← no extra energy in form of ATP/GTP is needed



5.5 Self-Splicing Introns, Detailed Structure



- small protein
  - not needed in vitro for reaction
  - does not directly participate in catalysis
  - helps intron RNA to fold (helps to shield some of the repulsive charge)
- present in tightly packed RNA

crystal structure (both group I and group II)
- tightly packed double-stranded RNA structures
- group II slightly bigger (!)

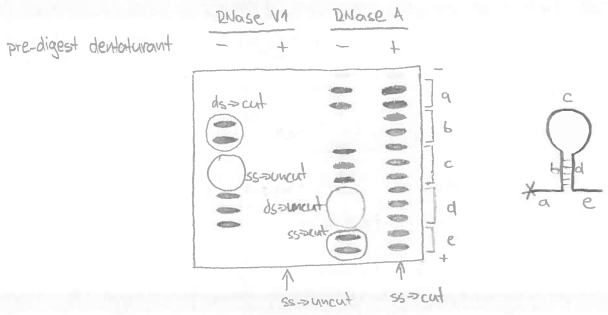
← ribozymes

group II self-splicing RNA

5.6. Assays for RNA Structure, RNase Digestion

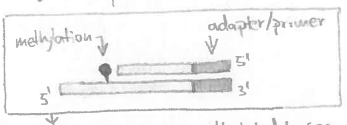
- RNase digestion: assay for RNA 2° structure (in vitro)
  - probing for regions of ss vs. ds RNA

RNase
- V1: cleaves dsRNA (but cuts only one strand)
- A: cleaves 3' of ss C/U
- T1: cleaves 3' of ss G



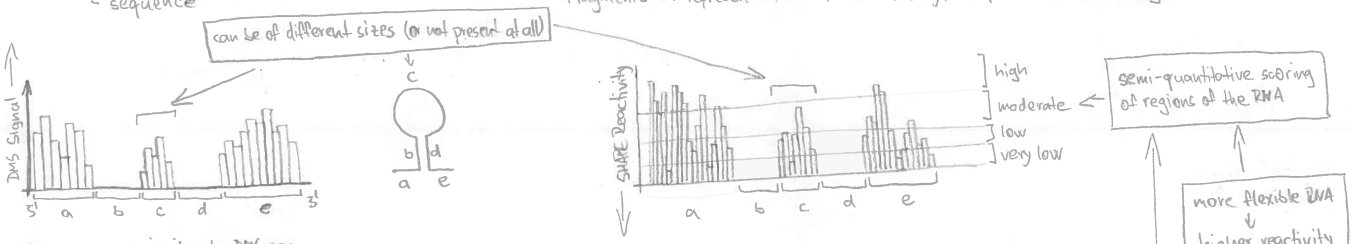
5.7. Assays for RNA Structure, DMS-seq and SHAPE-seq

- DMS-seq: assay for RNA 2° structure (in vivo/in vitro) ← different readouts for in vivo/in vitro signifies improperly folded RNA in vitro
  - probing for regions of ss vs. ds RNA
  - DMS (dimethyl sulfide)
    - methylating agent
    - methylates ss regions (ds regions well protected, since methylation occurs in part of RNA that forms H-bonds in ds regions)
    - methylates only A/C
  - process
    - treat RNA w/ DMS (each RNA is methylated at few positions: ~1-2)
    - fragment methylated RNAs
    - size-select fragments (60-70 bp long)
    - ligate an adapter, anneal primer to the adapter & reverse transcriptase – does NOT replicate over methylated bases
    - size select for shorter ⇑ fragments (e.g. 25-45 bp long) ⇒ weed out all products of unmethylated templates
    - sequence



⇒ fragments will represent distribution of methylated positions ⇒ DMS signal

# of 3' ends of size-selected fragments that align to given position



- SHAPE-seq – very similar to DMS-seq
  - SHAPE: Selective 2'Hydroxyl Acylation and Primer Extension
  - reagent: 1M7
    - adds larger modification to 2'OH ← reverse transcriptase does NOT replicate over the 2'OH modification
    - modification NOT restricted to particular bases ← method superior to DMS, which is restricted to A/C
- both assays (DMS-seq & SHAPE-seq) provide the same resolution

5.8. Assays for RNA Structure, Computational Prediction

- energy minimization algorithm
  - optimize the lowest energy state of folded RNA by
    - maximizing regions of base-pairing
    - minimizing too small loops
- covariation – if the purpose of RNA region is to form ds-structure, certain mutations will not be selected against (e.g. C-G vs G-C)

will be present in homologous RNAs

5.9. Splicing Mechanism

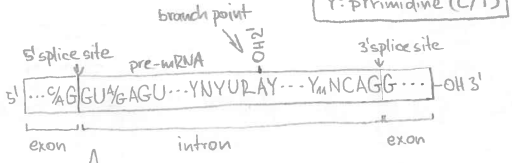
- pre-mRNA splicing
  - very common in some organisms (especially, but not limited to, higher eukaryotes)
  - similar to group II intron splicing (see ch. 5.4., pg. 5.1)

	group II self-splicing	pre-mRNA splicing
catalyzed by	intron RNA	RNAs and proteins
intron structure	long, highly structured	less information, less structure
splicing reactions	one rxn, catalyzed by intron	many rxns per pre-mRNA, catalyzed by splicing machinery

two-step transesterification

e.g. ~7 introns on average in human pre-mRNA

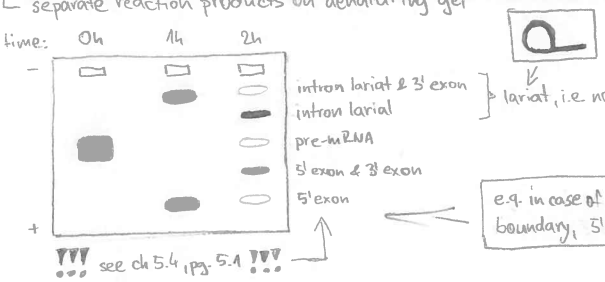
R: purine (A/G)
Y: pYrimidine (C/T)



- usually, multiple introns are present
- all introns/exons have the same conserved sequences
  - ⇒ correct introns/exons need to be brought together by the splicing machinery
    - usually neighboring ones, or
    - alternative splicing

5.10. Discovery of Splicing Mechanism, Cis Components

- in-vitro splicing assay
- process
  - synthetize body-labeled pre-mRNA
    - e.g. α ^{32}P-UTP (UTP w/ ^{32}P in α-phosphate)
      - → only RNA will be labeled (U)
      - → label incorporated into RNA backbone (αP)
  - incubate pre-mRNA in splicing-competent nuclear extract
    - → splicing happens in nucleus!!!
    - can be made from cells from different organisms
      - higher eukaryotes: usually do work on each others RNAs (e.g. chicken→human, Hela cells→alligator, ...)
      - yeast: may not properly recognize sites on higher euk. pre-mRNA
  - separate reaction products on denaturing gel



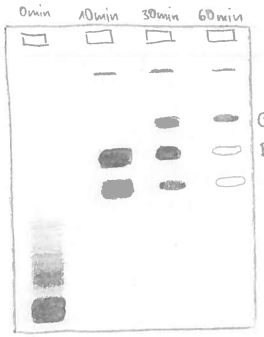
lariat, i.e. non-linear structure
- travels slower in the gel
- position will not correspond to molecular weight

e.g. in case of mutation in intron conserved sequence that prevents 5'exon from recognizing the lariat/3'exon boundary, 5'exon/3'exon mature mRNA will not be formed and gel will look the same for 1h and 2h lanes

!!! see ch 5.4, pg. 5.1 !!!

5.11. Discovery of Splicing Mechanism, Trans Components

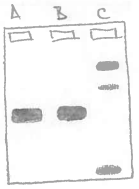
- same in-vitro splicing assay ↑ (ch 5.10)
- process – use native gel instead of denaturing gel (otherwise the same)



need to fractionate components of these complexes in order to study them more closely

one of the most gentle ways of separating molecules from each other (advantageous in working w/ big complexes)

- variation
  - separate reaction products in an ultracentrifuge using a gradient (glycerol, sucrose)
  - run products from the gradient on a denaturing gel



can observe how many steps of splicing have occured

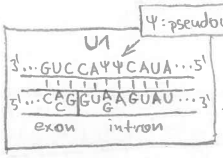
process
- synthesize body-labeled pre-mRNA substrate
- incubate in splicing-competent extract
- separate rxn products on gradients (ultracentrifuge)
- RNA-protein complexes separate on gradients based on size
- collect gradient fractions & monitor radioactivity
- isolate RNA from fractions
- separate RNA species on denaturing gel

5.12 Spliceosome Cycle, Pt. I

- spliceosome component types
  - snRNPs (Small Nuclear RiboNucleo Proteins)
    - RNA molecule (~100 nt long) bound w/ proteins that help it to do its function
  - free proteins
- U1: experiments
  - is U1-pre-mRNA complementarity important for function?
    - experiment: mutate pre-mRNA & observe (change in) function
      - → U1
  - is U1 important factor for splicing?
    - does removal of U1 prevent splicing?
      - experiment: deplete splicing extract of U1 using antibody
    - is U1 identifiable in splicing complex?
      - experiment: in-vitro splicing assay, native gel, western blot

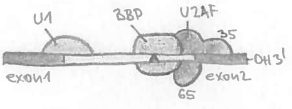
modified post-trn

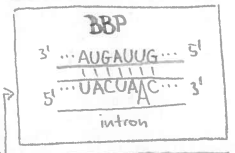
Ψ: pseudouridine



E complex
spliceosome cycle

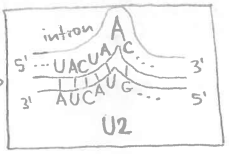
E: Early





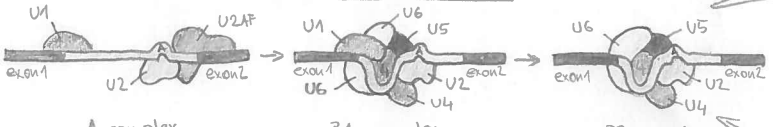
- BBP: Branchpoint Binding Protein
- U2AF: U2 Associated Factor
  - two subunits: 35, 65
  - molecular weights

mutually exclusive



role of ATP: specificity and order (regulation)

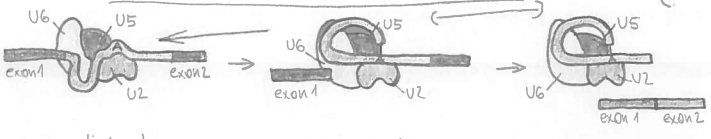
5.13 Spliceosome Cycle, Pt. II



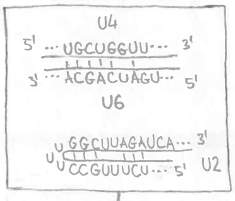
B1→B2: U1 removed by DEAD-box proteins
- low processivity
- ATPases (helicases)
- active site: Asp(D), Glu(E), Ala(A), Asp(D)

e.g.

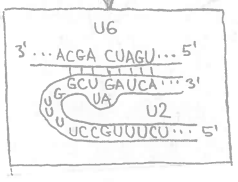
5.14. Spliceosome Cycle, Pt. III



break interaction b/w U4/U6
U6 interacts w/ U2



recycling
- debranch the lariat and recycle its RNA
- reconstitute U6/U4/U5 snRNP
  - assisted by proteins



key component of catalytic reaction

- snRNPs
  - U1: recognizes & binds to 5' splice site (included in complexes A, B)
  - U2: recognizes & binds the branch point sequence (included in complexes A, B, C)
  - U4: chaperon for U6
  - U5: structural scaffold
  - U6: contains active site