Studying the RNA Transcript of a Gene

Transcript analysis by northern hybridization:

- **Northern hybridization**, the RNA equivalent of Southern hybridization, is used to measure the length of a transcript.
- An RNA extract is electrophoresed in an agarose gel, using a denaturing electrophoresis buffer (e.g., one containing formaldehyde) to ensure that the RNAs do not form inter- or intramolecular base pairs, as base pairing would affect the rate at which the molecules migrate through the gel.
- After electrophoresis, the gel is blotted onto a nylon or nitrocellulose membrane, and hybridized with a labeled probe. If the probe is a cloned gene, the band that appears in the autoradiograph is the transcript of that gene.
- The size of the transcript can be determined from its position within the gel, and if RNA from different tissues is run in different lanes of the gel, then the possibility that the gene is differentially expressed can be examined.

Studying the RNA Transcript of a Gene

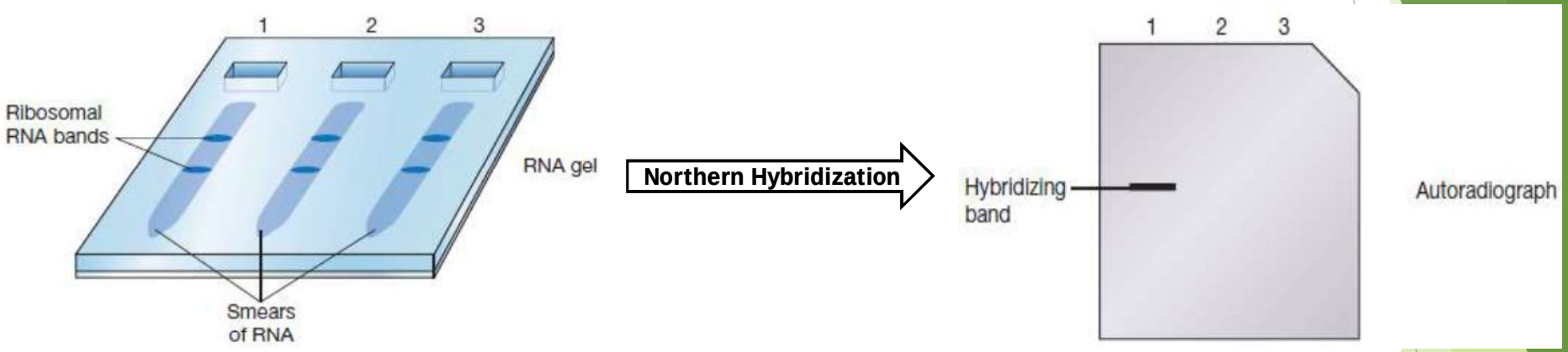

FIG: Northern hybridization. Three RNA extracts from different tissues have been electrophoresed in an agarose gel. The extracts are made up of many RNAs of different lengths so each gives a smear of RNA, but two distinct bands are seen, one for each of the abundant ribosomal RNAs. The sizes of these rRNAs are known (e.g. 4718 and 1874 nucleotides in mammals), so they can be used as internal size markers. The gel is transferred to a membrane, probed with a cloned gene, and the results visualized, for example by autoradiography if the probe has been radioactively labeled. Only lane 1 gives a band, showing that the cloned gene is expressed only in the tissue from which this RNA extract was obtained.

Studying the RNA Transcript of a Gene

Transcript analysis by DNA-mRNA hybridization:

- **Nucleic acid hybridization** occurs just as readily between complementary DNA and RNA strands as between single-stranded DNA molecules. If a hybrid is formed between a DNA strand, containing a gene, and its mRNA, then the boundaries between the double- and single-stranded regions will mark the start and end points of the mRNA.
- Introns, which are present in the DNA but not in the mRNA, will "loop out" as additional single-stranded regions.
- Now consider the result of treating the DNA–RNA hybrid with a single-strand specific nuclease such as S1. S1 nuclease degrades single-stranded DNA or RNA polynucleotides, including single-stranded regions at the ends of predominantly double stranded molecules, but has no effect on double-stranded DNA or on DNA–RNA hybrids.

Studying the RNA Transcript of a Gene

- S1 nuclease will therefore digest the non-hybridized single-stranded DNA regions at each end of the DNA–RNA hybrid, along with any looped-out introns. The single-stranded DNA fragments protected from S1 nuclease digestion can be recovered if the RNA strand is degraded by treatment with alkali.

Limitation of the method:

- Although the sizes of the protected DNA fragments could be measured by gel electrophoresis, this does not allow their order or relative positions in the DNA sequence to be determined.

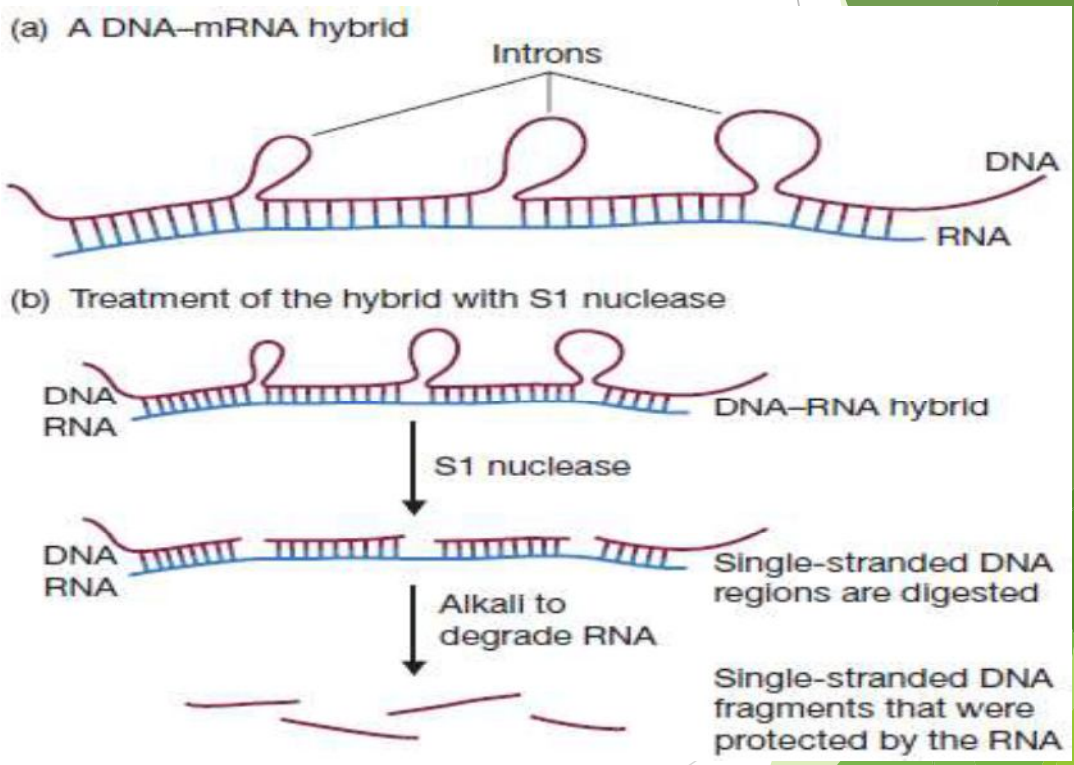

FIG: A DNA-mRNA hybrid and the effect of treating this hybrid with a single-strand-specific nuclease such as S1.

Studying the RNA Transcript of a Gene

Locating a transcription start point by S1 nuclease mapping:

- A few subtle modifications to the previous technique allow the precise start and end points of the transcript and of any introns it contains to be mapped onto the DNA sequence. An example of the way in which **S1 nuclease mapping** is used to locate the start point of a transcript is shown in the figure.
- Here, a *Sau*3A fragment that contains 100 bp of coding region, along with 300 bp of the leader sequence preceding the gene, has been cloned into an M13 vector and obtained as a single-stranded molecule. A sample of the RNA transcript is added and allowed to anneal to the DNA molecule.

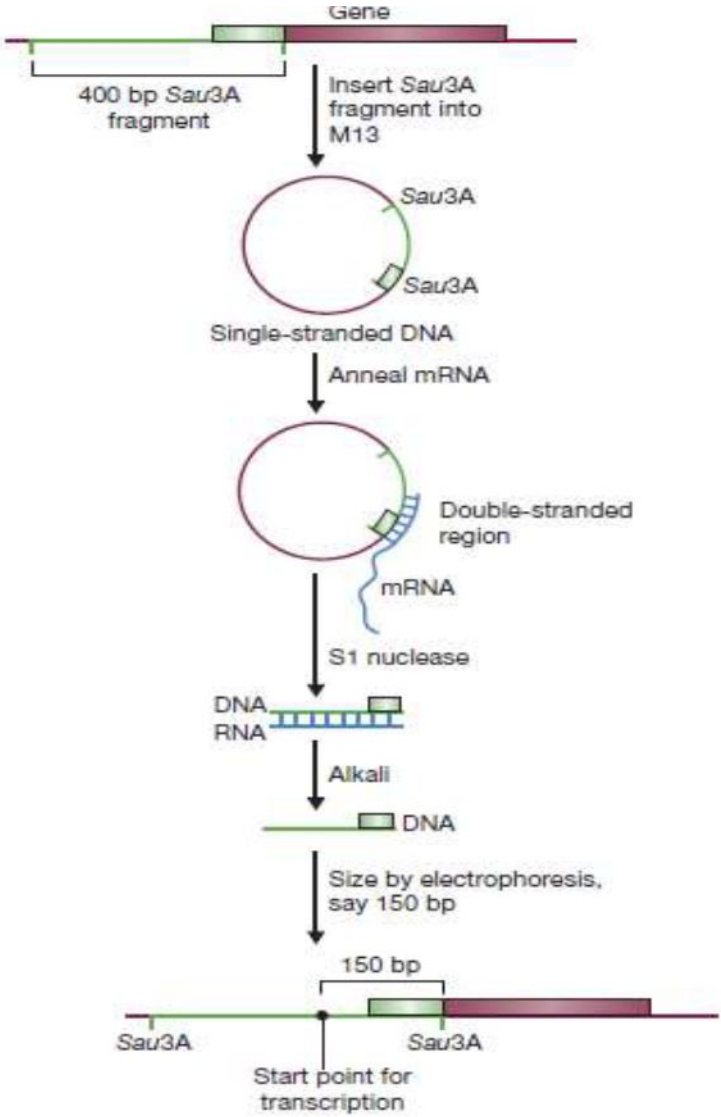

FIG: Locating a transcription start point by S1 nuclease mapping.

Studying the RNA Transcript of a Gene

- The DNA molecule is still primarily single-stranded but now has a small region protected by the RNA transcript. All but this protected region is digested by S1 nuclease and the RNA is degraded with alkali, leaving a short single-stranded DNA fragment.
- If these manipulations are examined closely it will become clear that the size of this single stranded fragment corresponds to the distance between the transcription start point and the right-hand *Sau*3A site. The size of the single-stranded fragment is therefore determined by gel electrophoresis and this information is used to locate the transcription start point on the DNA sequence. Exactly the same strategy could locate the end point of transcription and the junction points between introns and exons.

Studying the RNA Transcript of a Gene

Transcript analysis by primer extension:

- **Primer extension—**is less adaptable, because it can only identify the 5′ end of an RNA. It is, nonetheless, an important technique that is frequently used to check the results of S1 analyses.
- Primer extension can only be used if at least part of the sequence of the transcript is known. This is because a short oligonucleotide primer must be annealed to the RNA at a known position, ideally within 100–200 nucleotides of the 5′ end of the transcript.
- Once annealed, the primer is extended by reverse transcriptase . This is a cDNA synthesis reaction, but one that is very likely to proceed to completion as only a short segment of RNA has to be copied.

Studying the RNA Transcript of a Gene

- The 3′ end of the newly synthesized strand of DNA will therefore correspond with the 5′ terminus of the transcript.
- Locating the position of this terminus on the DNA sequence is achieved simply by determining the length of the single-stranded DNA molecule and correlating this information with the annealing position of the primer.

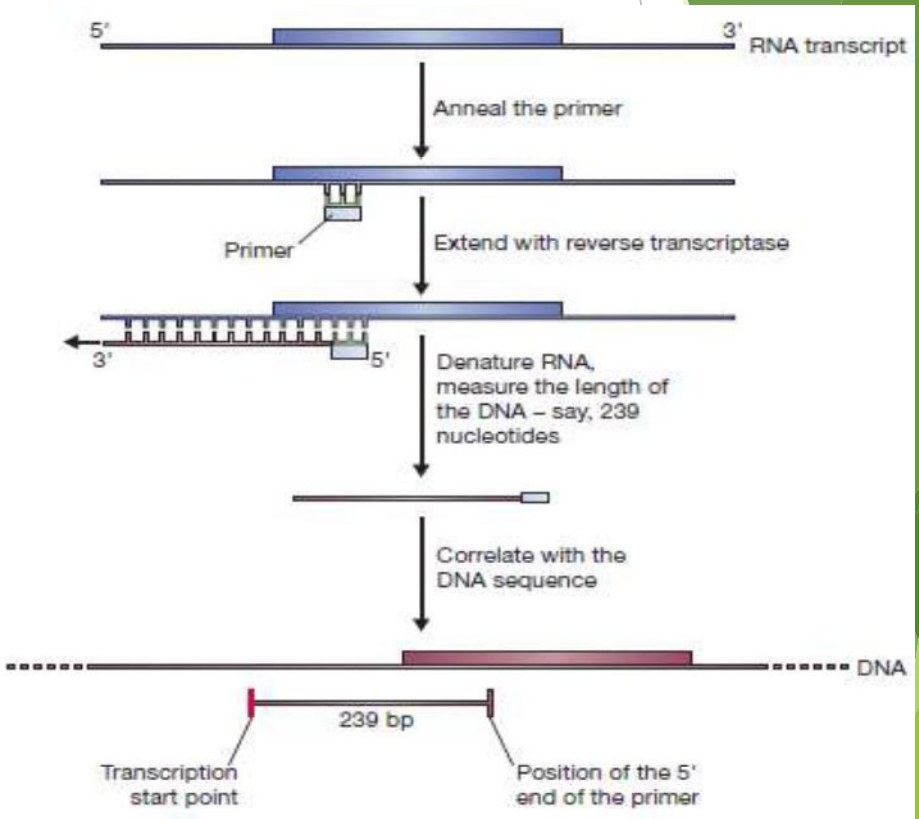

FIG: Locating a transcription start point by primer extension.

Studying the RNA Transcript of a Gene

Transcript analysis by PCR:

- **Rapid amplification of cDNA ends (RACE)** can be used to identify the 5′ and 3′ termini of RNA molecules and so, like S1 analysis, can be used to map the ends of transcripts. There are several variations to the RACE method. Here we will consider how the 5′ end of an RNA molecule can be mapped.
- This procedure uses a primer that is specific for an internal region of the RNA molecule. The primer attaches to the RNA and directs the first, reverse transcriptase catalyzed, stage of the process, during which a single-stranded cDNA is made. As in the primer extension method, the 3′ end of the cDNA corresponds with the 5′ end of the RNA.

Studying the RNA Transcript of a Gene

- Terminal deoxynucleotidyl transferase is now used to attach a series of A nucleotides to the 3′ end of the cDNA, forming the priming site for a second PCR primer, which is made up entirely of Ts and hence anneals to the poly(A) tail created by terminal transferase. Now the standard PCR begins, first converting the single-stranded cDNA into a double-stranded molecule, and then amplifying this molecule as the PCR proceeds. The PCR product is then sequenced to reveal the precise position of the start of the transcript.

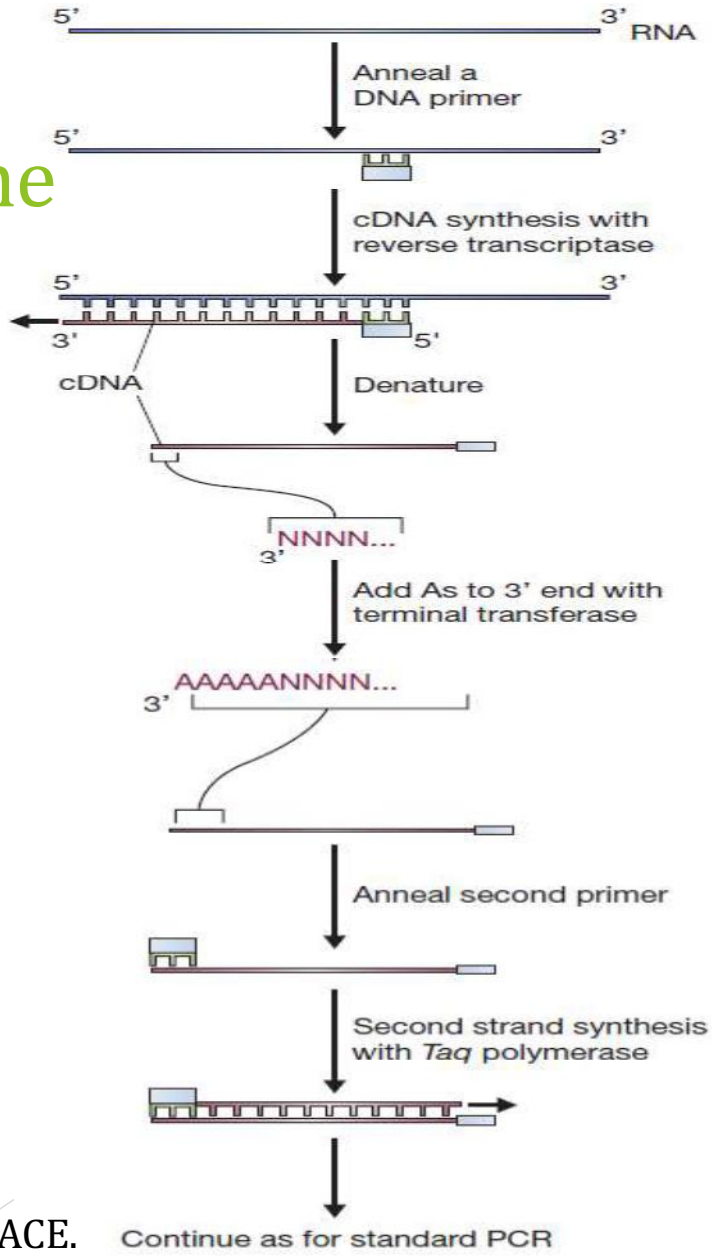

FIG: Identification of 5′ terminus by RACE.