8

RNA: Transcription and Processing

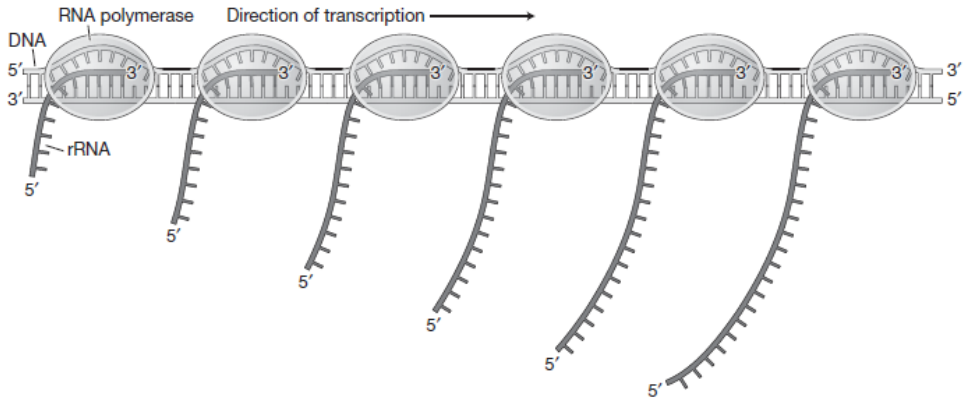
WORKING WITH THE FIGURES

1. In Figure 8-3, why are the arrows for genes 1 and 2 pointing in opposite directions?

Answer: The arrows for genes 1 and 2 indicate the direction of transcription, which is always 5′ to 3′. The two genes are transcribed from opposite DNA strands, which are antiparallel, so the genes must be transcribed in opposite directions to maintain the 5′ to 3′ direction of transcription.

2. In Figure 8-5, draw the "one gene" at much higher resolution with the following components: DNA, RNA polymerase(s), RNA(s).

Answer: At the higher resolution, the feathery structures become RNA transcripts, with the longer transcripts occurring nearer the termination of the gene. The RNA in this drawing has been straightened out to illustrate the progressively longer transcripts.



3. In Figure 8-6, describe where the gene promoter is located.

Answer: The promoter is located to the left (upstream) of the 3′ end of the template strand. From this sequence it cannot be determined how far the promoter would be from the 5′ end of the mRNA.

4. In Figure 8-9b, write a sequence that could form the hairpin loop structure.

Answer: Any sequence that contains inverted complementary regions separated by a noncomplementary one would form a hairpin. One sequence would be:

ACG**CAAGCUUAC**CGAUUAUU**GUAAGCUUG**AAG

The two bold-faced sequences would pair and form a hairpin. The intervening non-bold sequence would be the loop.

5. How do you know that the events in Figure 8-13 are occurring in the nucleus?

Answer: The figure shows a double-stranded DNA molecule from which RNA is being transcribed. This process only occurs in the nucleus.

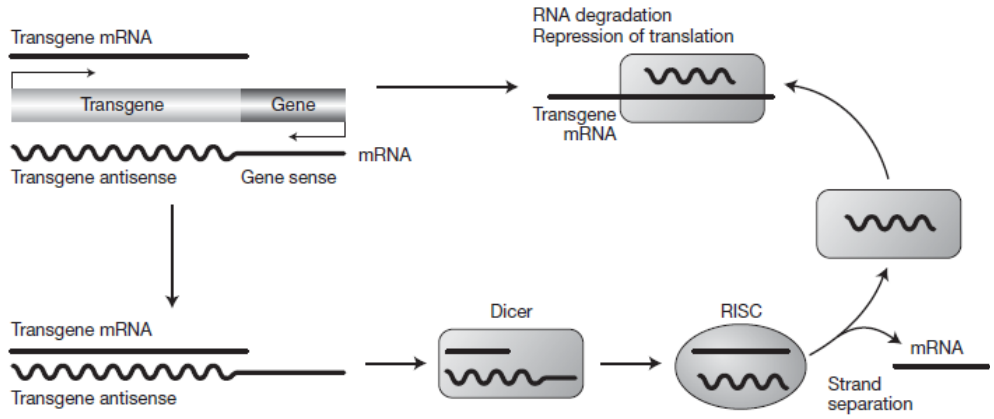
6. In Figure 8-15, what do you think would be the effect of a G to A mutation in the first G residue of the intron?

Answer: A mutation of G to A would alter the U1 SNP binding site and prevent formation of the spliceosome. This would prevent splicing of the intron.

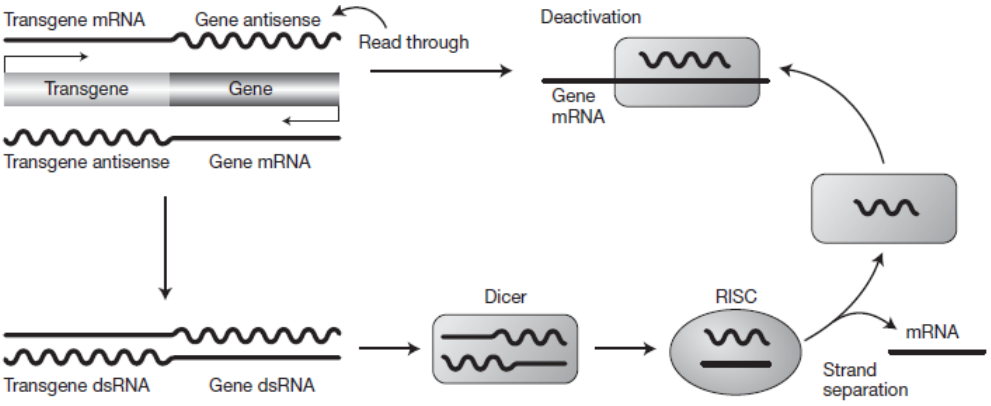
7. In Figure 8-23, show how the double-stranded RNA is able to silence the transgene. What would have to happen for the transgene to also silence the flanking cellular gene (in yellow)?

Answer:

a. The double-stranded RNA formed from the sense and antisense transgene transcripts would be processed by Dicer, then bind to RISC. RISC would separate the dsRNA to produce an antisense RNA/RISC complex. The RISC/RNA complex would bind to the transgene mRNA and deactivate it.



b. The flanking gene would be silenced if transcription of the transgene continued into the flanking gene. This would produce an antisense transcript of the flanking gene, leading to the formation of dsRNA and activation of Dicer and RISC.



BASIC PROBLEMS

8. The two strands of λ phage DNA differ from each other in their GC content. Owing to this property, they can be separated in an alkaline cesium chloride gradient (the alkalinity denatures the double helix). When RNA synthesized by l phage is isolated from infected cells, it is found to form DNA–RNA hybrids with both strands of λ DNA. What does this finding tell you? Formulate some testable predictions.

Answer: Because RNA can hybridize to both strands, the RNA must be transcribed from both strands. This does not mean, however, that both strands are used as a template *within each gene.* The expectation is that only one strand is used within a gene but that different genes are transcribed in different directions along the DNA. The most direct test would be to purify a specific RNA coding for a specific protein and then hybridize it to the λ genome. Only one strand should hybridize to the purified RNA.

9. In both prokaryotes and eukaryotes, describe what else is happening to the RNA while RNA polymerase is synthesizing a transcript from the DNA template.

Answer: In prokaryotes, translation is beginning at the 5′ end while the 3′ end is still being transcribed. In eukaryotes, processing (capping, splicing) is occurring at the 5′ end while the 3′ end is still being transcribed.

10. List three examples of proteins that act on nucleic acids.

Answer: There are many examples of proteins that act on nucleic acids, but some mentioned in this chapter are RNA polymerase, GTFs (general transcription factors), σ (sigma factor), rho, TBP (TATA binding protein), and snRNPs (a combination of proteins and snRNAs).

11. What is the primary function of the sigma factor? Is there a protein in eukaryotes analogous to the sigma factor?

Answer: Sigma factor, as part of the RNA polyermase holoenzyme, recognizes and binds to the –35 and –10 regions of bacterial promoters. It positions the holoenzyme to correctly initiate transcription at the start site. In eukaryotes, TBP (TATA binding protein) and other GTFs (general transcription factors) have an analagous function.

12. You have identified a mutation in yeast, a unicellular eukaryote, that prevents the capping of the 5′ end of the RNA transcript. However, much to your surprise, all the enzymes required for capping are normal. You determine that the mutation is, instead, in one of the subunits of RNA polymerase II. Which subunit is mutant and how does this mutation result in failure to add a cap to yeast RNA?

Answer: The CTD (carboxy tail domain) of the ß subunit of RNA polymerase II contains binding sites for enzymes and other proteins that are required for RNA processing and is located near the site where nascent RNA emerges. If mutations in this subunit prevent the correct binding and/or localization of the proteins necessary for capping, then this modification will not occur even though all the required enzymes are normal.

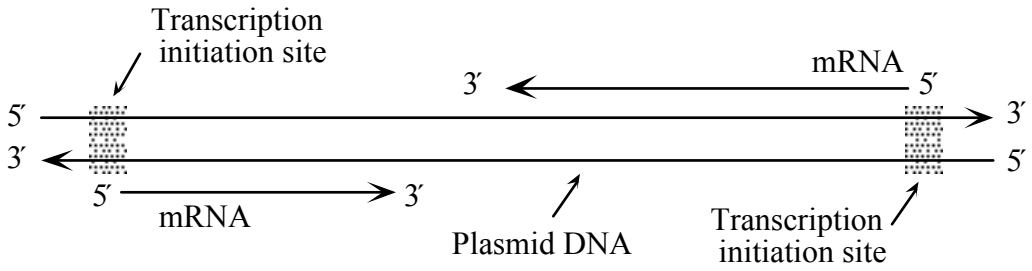
13. Why is RNA produced only from the template DNA strand and not from both strands?

Answer: For a given gene, only one strand of DNA is transcribed. This strand (called the template) will be complementary to the RNA and also to the other strand (called the nontemplate or coding strand). Consequently, the nucleotide sequence of the RNA must be the same as that of the nontemplate strand of the DNA (except that the Ts are instead Us). Ultimately, it is the nucleotide sequence that gives the RNA its function. Transcription of both strands would give two complementary RNAs that would code for completely different polypeptides. Also, double-stranded RNA initiates cellular processes that lead to its degradation.

14. A linear plasmid contains only two genes, which are transcribed in opposite directions, each one from the end, toward the center of the plasmid. Draw diagrams of

a. the plasmid DNA, showing the 5′ and 3′ ends of the nucleotide strands.
b. the template strand for each gene.
c. the positions of the transcription-initiation site.
d. the transcripts, showing the 5′ and 3′ ends.

Answer:
a., b., c., and d.



15. Are there similarities between the DNA replication bubbles and the transcription bubbles found in eukaryotes? Explain.

Answer: Yes. Both replication and transcription is performed by large, multi-subunit molecular machines (the replisome and RNA polymerase II, respectively) and both require helicase activity at the fork of the bubble. However, transcription proceeds in only one direction and only one DNA strand is copied.

16. Which of the following statements are true about eukaryotic mRNA?

a. The sigma factor is essential for the correct initiation of transcription.
b. Processing of the nascent mRNA may begin before its transcription is complete.

c. Processing takes place in the cytoplasm.
d. Termination is accomplished by the use of a hairpin loop or the use of the rho factor.
e. Many RNAs can be transcribed simultaneously from one DNA template.

Answer:
a. False. Sigma factor is required in prokaryotes, not eukaryotes.
b. True. Processing begins at the 5′ end, while the 3′ end is still being synthesized.
c. False. Processing occurs in the nucleus and only mature RNA is transported out to the cytoplasm.
d. False. Hairpin loops or rho factor (in conjunction with the *rut* site) is used to terminate transcription in prokaryotes. In eukaryotes the conserved sequences AAUAAA or AUUAAA, near the 3′ end of the transcript, are recognized by an enzyme that cuts off the end of the RNA approximately 20 bases downstream.
e. True. Multiple RNA polymerases may transcribe the same template simultaneously.

17. A researcher was mutating prokaryotic cells by inserting segments of DNA. In this way, she made the following mutation:

Original TTGACAT 15 to 17 bp TATAAT
Mutant TATAAT 15 to 17 bp TTGACAT

a. What does this sequence represent?
b. What do you predict will be the effect of such a mutation? Explain.

Answer:
a. The original sequence represents the –35 and –10 consensus sequences (with the correct number of intervening spaces) of a bacterial promoter. Sigma factor, as part of the RNA polymerase holoenzyme, recognizes and binds to these sequences.
b. The mutated (transposed) sequences would not be a binding site for sigma factor. The two regions are not in the correct orientation to each other and therefore would not be recognized as a promoter.

18. You will learn more about genetic engineering in Chapter 10, but for now, put on your genetic engineer's cap and try to solve this problem. E. coli is widely used in laboratories to produce proteins from other organisms.

a. You have isolated a yeast gene that encodes a metabolic enzyme and want to produce this enzyme in *E. coli*. You suspect that the yeast promoter will not work in *E. coli*. Why?

b. After replacing the yeast promoter with an *E. coli* promoter, you are pleased to detect RNA from the yeast gene but are confused because the RNA is almost twice the length of the mRNA from this gene isolated from yeast. Explain why this result might have occurred.

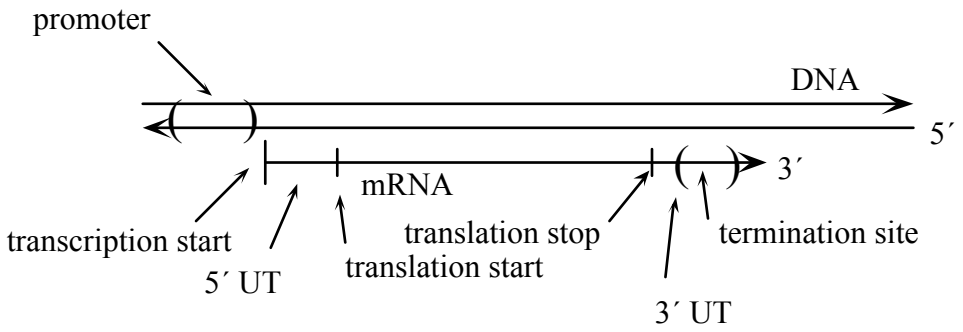
Answer:

a. The promoters of eukaryotes and prokaryotes do not have the same conserved sequences. In yeast, the promoter would have the required TATA box located about –30, whereas bacteria would have conserved sequences at –35 and –10 that would interact with sigma factor as part of the RNA polymerase holoenzyme.

b. There are two possible reasons that the mRNA is longer than expected. First, many eukaryotic genes contain introns, and bacteria would not have the splicing machinery necessary for their removal. Second, termination of transcription is not the same in bacteria and yeast; the sequences necessary for correct termination in *E. coli* would not be expected in the yeast gene.

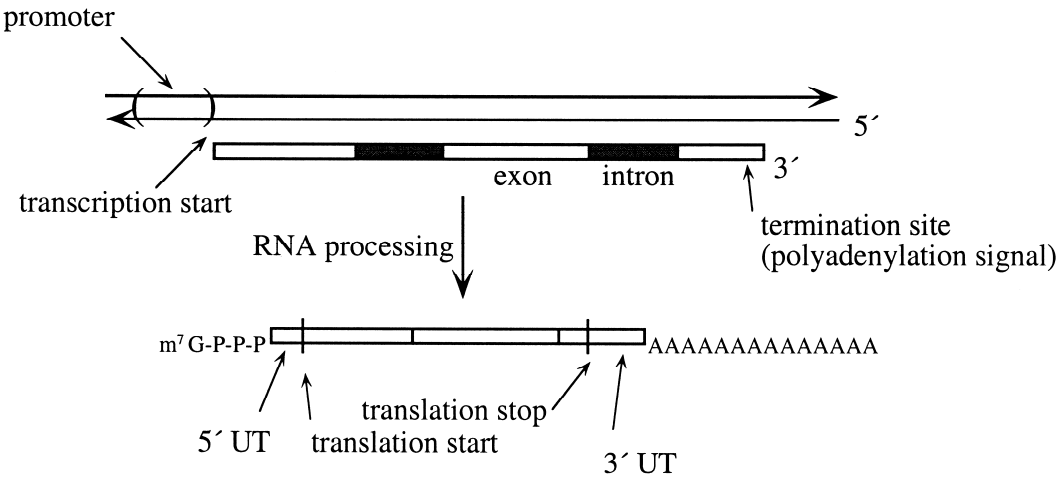
19. Draw a prokaryotic gene and its RNA product. Be sure to include the promoter, transcription start site, transcription termination site, untranslated regions, and labeled 5′ and 3′ ends.

Answer:



20. Draw a two-intron eukaryotic gene and its pre-mRNA and mRNA products. Be sure to include all the features of the prokaryotic gene included in your answer to Problem 19, plus the processing events required to produce the mRNA.

Answer:



21. A certain *Drosophila* protein-encoding gene has one intron. If a large sample of null alleles of this gene is examined, will any of the mutant sites be expected

a. in the exons?
b. in the intron?
c. in the promoter?
d. in the intron–exon boundary?

Answer:
a. Yes. The exons encode the protein, so null mutations would be expected to map within exons.
b. Possibly. There are sequences near the boundaries of and within introns that are necessary for correct splicing. If these are altered by mutation, correct splicing will be disrupted. Although transcribed, it is likely that translation will not occur.
c. Yes. If the promoter is deleted or altered such that GTFs cannot bind, transcription will be disrupted.
d. Yes. There are sequences near the boundaries of and within introns that are necessary for correct splicing.

22. What are self-splicing introns and why does their existence support the theory that RNA evolved before protein?

Answer: Self-splicing introns are capable of excising themselves from a primary transcript without the need of additional enzymes or energy source. They are one of many examples of RNA molecules that are catalytic, and for this property, they are also known as ribozymes. With this additional function, RNA is the only known biological molecule to encode genetic information and catalyze biological reactions. In simplest terms, it is possible that life began

with an RNA molecule, or group of molecules, that evolved the ability to self-replicate.

23. Antibiotics are drugs that selectively kill bacteria without harming animals. Many antibiotics act by selectively binding to certain proteins that are critical for bacterial function. Explain why some of the most successful antibiotics target bacterial RNA polymerase.

Answer: Antibiotics need to selectively target bacterial structures and functions that are essential for life but unique or sufficiently different from the equivalent structure and functions of their animal hosts. Bacterial RNA polymerase fits these criteria as its function is obviously essential, yet its structure is sufficiently different from the several eukaryotic RNA polymerases. These differences make it possible to develop drugs that specifically bind bacterial RNA polymerase but have little or no affinity for eukaryotic RNA polymerases.

CHALLENGING PROBLEMS

24. The following data represent the base compositions of double-stranded DNA from two different bacterial species and their RNA products obtained in experiments conducted in vitro:

Species	(A + T)/ (G + C)	(A + U)/ (G + C)	(A + G)/ (U + C)
Bacillus subtilis	1.36	1.30	1.02
E. coli	1.00	0.98	0.80

a. From these data, can you determine whether the RNA of these species is copied from a single strand or from both strands of the DNA? How? Drawing a diagram will make it easier to solve this problem.
b. Explain how you can tell whether the RNA itself is single stranded or double stranded.

(Problem 24 is reprinted with the permission of Macmillan Publishing Co., Inc., from M. Strickberger, *Genetics*. Copyright 1968, Monroe W. Strickberger.)

Answer:
a. The data cannot indicate whether one or both strands are used for transcription. You do not know how much of the DNA is transcribed nor which regions of DNA are transcribed. Only when the purine/pyrimidine ratio is not unity can you deduce that only one strand is used as template.

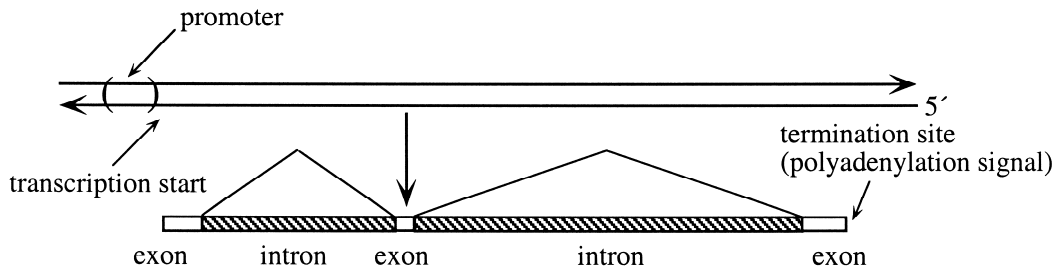
b. If the RNA is double-stranded, the percentage of purines (A + G) would equal the percentage of pyrimidines (U + C) and the (A + G)/(U + C) ratio would be 1.0. This is clearly not the case for *E. coli*, which has a ratio of

0.80. The ratio for *B. subtilis* is 1.02. This is consistent with the RNA being double-stranded but does not rule out single-stranded if there are an equal number of purines and pyrimidines in the strand.

25. A human gene was initially identified as having three exons and two introns. The exons are 456, 224, and 524 bp, whereas the introns are 2.3 kb and 4.6 kb.

a. Draw this gene, showing the promoter, introns, exons, and transcription start and stop sites.
b. Surprisingly, this gene is found to encode not one but two mRNAs that have only 224 nucleotides in common. The original mRNA is 1204 nucleotides, and the new mRNA is 2524 nucleotides. Use your drawing to show how this one region of DNA can encode these two transcripts.

Answer:
a.



b. Alternative splicing of the primary transcript would result in mRNAs that were only partially identical. In this case, the two transcripts share 224 nucleotides in common. As this is the exact length of the second exon, one possible solution to this problem is that this exon is shared by the two alternatively-spliced mRNAs. The second transcript also contains 2.3 kb of sequence not found in the first. Perhaps what was considered the first intron is actually also part of the second transcript as that would result in the 2524 nucleotides stated as this transcript's length. Of course, other combinations of alternative splicing would also fit the data.

26. While working in your laboratory, you isolate an mRNA from *C. elegans* that you suspect is essential for embryos to develop successfully. With the assumption that you are able to turn mRNA into double-stranded RNA, design an experiment to test your hypothesis.

Answer: Double-stranded RNA, composed of a sense strand and a complementary antisense strand, can be used in *C. elegans* (and likely all organisms) to selectively prevent the synthesis of the encoded gene product (a discovery awarded the 2006 Nobel Prize in Medicine). This process, called gene silencing, blocks the synthesis of the encoded protein from the endogenous gene

and is thus equivalent to "knocking out" the gene. To test whether a specific mRNA encodes an essential embryonic protein, eggs or very early embryos should be injected with the double-stranded RNA produced from your mRNA, thus activating the RNAi pathway. The effects of knocking out the specific gene product can then be followed by observing what happens in these, versus control, embryos. If the encoded protein is essential, embryonic development should be perturbed when your gene is silenced.

27. Glyphosate is an herbicide used to kill weeds. It is the main component of a product made by the Monsanto Company called Roundup. Glyphosate kills plants by inhibiting an enzyme in the shikimate pathway called EPSPS. This herbicide is considered safe because animals do not have the shikimate pathway. To sell even more of their herbicide, Monsanto commissioned its plant geneticists to engineer several crop plants, including corn, to be resistant to glyphosate. To do so, the scientists had to introduce an EPSPS enzyme that was resistant to inhibition by glyphosate into crop plants and then test the transformed plants for resistance to the herbicide.

Imagine that you are one of these scientists and that you have managed to successfully introduce the resistant EPSPS gene into the corn chromosomes. You find that some of the transgenic plants are resistant to the herbicide, whereas others are not. Your supervisor is very upset and demands an explanation of why some of the plants are not resistant even though they have the transgene in their chromosomes. Draw a picture to help him understand.

Answer: Transgene silencing is a common phenomenon in plants. Silencing may occur at the transcriptional or post-transcriptional level. Since you cannot control where transgenes insert, some may insert into transcriptionally inactive parts of the genome. Post-transcriptional silencing may be the result of activation of the RNAi pathway due to the misexpression of both strands of your transgene. See Figure 8-22 in the companion text as one example of how this can happen. In these cases, the RNAi pathway will be activated and the EPSPS gene product will be silenced.

28. Many human cancers result when a normal gene mutates and leads to uncontrolled growth (a tumor). Genes that cause cancer when they mutate are called oncogenes. Chemotherapy is effective against many tumors because it targets rapidly dividing cells and kills them. Unfortunately, chemotherapy has many side effects, such as hair loss or nausea, because it also kills many of our normal cells that are rapidly dividing, such as those in the hair follicles or stomach lining.

Many scientists and large pharmaceutical companies are excited about the prospects of exploiting the RNAi pathway to selectively inhibit oncogenes in life-threatening tumors. Explain in very general terms how gene-silencing

therapy might work to treat cancer and why this type of therapy would have fewer side effects than chemotherapy.

Answer: RNAi has the potential to selectively prevent protein production from any targeted gene. Oncogenes are mutant versions of “normal” genes (called proto-oncogenes) and their altered gene products are partly or wholly responsible for causing cancer. In theory, it may be possible to design appropriate siRNA molecules (small interfering RNAs) that specifically silence the mutant oncogene product but do not silence the closely related proto-oncogene product. The latter is necessary to prevent serious side effects as the products of proto-oncogenes are essential for normal cellular function.