C1. A. tRNA genes encode tRNA molecules, and rRNA genes encode the rRNAs found in ribosomes. There are also genes for the RNAs found in snRNPs, etc.

B. The term *template strand* is still appropriate because one of the DNA strands is used as a template to make the RNA. The term *coding strand* is not appropriate because the RNA made from nonstructural genes does not code for a polypeptide sequence.

C. Yes.

C2. The formation of the open complex, and the release of sigma factor.

C3. A consensus sequence is the most common nucleotide sequence that is found within a group of related sequences. An example is the –35 and –10 consensus sequence found in bacterial promoters. At –35, it is TTGACA, but it can differ by one or two nucleotides and still function efficiently as a promoter. In the consensus sequences within bacterial promoters, the –35 site is primarily for recognition by sigma factor. The –10 site, also known as the Pribnow box, is the site where the DNA will begin to unwind to allow transcription to occur.

C4. GGCATTGTCA

C5. Mutations that make a sequence more like the consensus sequence are likely to be up promoter mutations while mutations that cause the promoter to deviate from the consensus sequence are likely to be down promoter mutations. Also, in the –10 region, AT pairs are favored over GC pairs, because the role of this region is to form the open complex. AT pairs are more easily separated because they form only two hydrogen bonds compared to GC pairs, which form three hydrogen bonds.

A. Up promoter

B. Down promoter

C. Up promoter

C6. The most highly conserved positions are the first, second, and sixth. In general, when promoter sequences are conserved, they are more likely to be important for binding. That explains why changes are not found at these positions; if a mutation altered a conserved position, the promoter would probably not work very well. By comparison, changes are occasionally tolerated at the fourth position and frequently at the third and fifth positions. The positions that tolerate changes are less important for binding by sigma factor.

C7. In Figure 12.5, each α–helix (labeled *2* and *3*) appears to occupy a region that is approximately one-half of a complete turn of the double helix. Since a complete turn of a DNA double helix involves 10 bp, each *a* helix appears to be bound to about 5 bp in the major groove of the DNA; 5 bp would span a linear distance of approximately 1.7 nm (as described in Chapter 9, Figure 9.17). If we divide 1.7 nm by 0.15 nm/amino acid, we obtain a value of 11.3 amino acids per helix. This would create about 3.15 turns of an *a* helix; there are three complete turns per helix, for a total of six complete turns for both helices 2 and 3.

C8. This will not affect transcription. However, it will affect translation by preventing the initiation of polypeptide synthesis.

C9. RNA polymerase holoenzyme consists of sigma factor plus the core enzyme, which is a tetramer, $a_2\beta\beta'$. The role of sigma factor is to recognize the promoter sequence. The *a* subunits are necessary for the assembly of the core enzyme and for loose DNA binding. The β and β' subunits are the portion that catalyze the covalent linkages between adjacent ribonucleotides.

C10. Sigma factor can slide along the major groove of the DNA. In this way, it is able to recognize base sequences that are exposed in the groove. When it encounters a promoter sequence, hydrogen bonding between the bases and the sigma factor protein can promote a tight and specific interaction.

C11. The mutation would alter the bases so that hydrogen bonding with the sigma factor would not occur, or would not occur in the correct way. From Figure 12.4, if we look at the –35 region, we may expect that changing the first two nucleotides to anything besides T might inhibit transcription, changing the third base to an A (perhaps), changing the fourth base to a G or T, or changing the last base to anything besides an A. We would think these substitutions may have an effect, because they would deviate from known functional bacterial promoters.

C12. DNA-G/RNA-C
DNA-C/RNA-G
DNA-A/RNA-U
DNA-T/RNA-A

The template strand is 3′–CCGTACGTAATGCCGTAGTGTGATCCCTAG–5′ and the coding strand is 5′–GGCATGCATTACGGCATCACACTAGGGATC–3′. The promoter would be to the left (in the 3′ direction) of the template strand.

C13. RNA polymerase slides down the DNA and forms an open complex as it goes. The open complex is a DNA bubble of about 17 nucleotide bp. Within the open complex, one of the DNA strands, the one running in the 3′ to 5′ direction, is used as template for RNA synthesis. This occurs as individual nucleotides hydrogen bond to the DNA template strand according to the rules described in conceptual question C12. As the RNA polymerase slides along, the DNA behind the open complex rewinds back into a double helix.

C14. Transcriptional termination occurs when the hydrogen bonding is broken between the DNA and the part of the newly made RNA transcript that is located in the open complex.

C15. In ρ-dependent termination, the p protein binds to the RNA transcript after the p site has been transcribed. Eventually, RNA polymerase will transcribe a termination stem-loop that will cause it to pause in the transcription process. As it is pausing, the p protein, which functions as a helicase, will catch up to RNA polymerase and knock it off the DNA. In p-independent transcription, there is no p protein. Again, the RNA polymerase transcribes a termination hairpin that causes it to pause. However, when it pauses, an AU-rich region is left base pairing in the open complex. Since this is holding on by fewer hydrogen bonds, it is rather unstable. Therefore, it tends to dissociate from the open complex and thereby end transcription.

C16. Helicase and p protein bind to a nucleic acid strand and travel in the 5′ to 3′ direction. When they encounter a double-stranded region, they break the hydrogen bonds between complementary strands. p protein is different from DNA helicase in that it moves along an RNA strand, while DNA helicase moves along a DNA strand. The purpose of DNA helicase function is to promote DNA replication while the purpose of p protein function is to promote transcriptional termination.

C17. RNA and DNA polymerase are similar in the following ways:

1. They both use a template strand.

2. They both synthesize in the 5′ to 3′ direction.

3. The chemistry of synthesis is very similar in that they use incoming triphosphates and make a phosphoester bond between the previous nucleotide and the incoming nucleotide.

4. They are both processive enzymes that slide along a template strand of DNA.

RNA and DNA polymerase are different in the following ways:

1. RNA polymerase makes an RNA strand while DNA polymerase makes a DNA strand.

2. RNA polymerase uses only one DNA strand as a template.

3. RNA polymerase recognizes a particular DNA sequence (i.e., promoters) and synthesizes only a defined region of RNA from the promoter to the terminator.

4. RNA polymerase cannot proofread.

5. RNA polymerase does not need a primer.

C18. A. Mutations that alter the AU-rich region by introducing Gs and Cs, and mutations that prevent the formation of the stem-loop (hairpin) structure.

B. Mutations that alter the termination sequence and mutations that alter the p recognition site.

C. Eventually, somewhere downstream from the gene, there would be found another transcriptional termination sequence and transcription would terminate there. This second termination sequence might be found randomly or it might be at the end of an adjacent gene.

C19. A. Ribosomal RNA (5.8S, 18S, and 28S)

B. All of the mRNA and certain snRNA genes

C. All of the tRNAs and the 5S rRNA

C20. Eukaryotic promoters are somewhat variable with regard to the pattern of sequence elements that may be found. In the case of structural genes that are transcribed by RNA polymerase II, it is common to have a TATA box, which is about 25 bp upstream from a transcriptional start site. The TATA box is important in the identification of the transcriptional start site and the assembly of RNA polymerase and various transcription factors. The transcriptional start site defines where transcription actually begins.

C21. A. RNA polymerase would not be bound to the TATA box.

B. TTIID contains the TATA-binding protein. If it were missing, RNA polymerase would not bind to the TATA box either.

C. The formation of the open complex would not take place.

C22. It is primarily an accessibility problem. When the DNA is tightly wound around histones, it becomes difficult for large proteins, like RNA polymerase and transcription factors, to recognize the correct base sequence in the DNA and to catalyze the movement of the open complex. It is thought that a partial or complete disruption of the nucleosome structure is necessary for transcription to occur.

C23. TFIID and TFIIB would play analogous roles. Sigma factor does two things. It recognizes the promoter (as does TFIID) and it recruits RNA polymerase to the promoter (as does TFIIB).

C24. Hydrogen bonding is usually the predominant type of interaction when proteins and DNA follow an assembly and disassembly process. In addition, ionic bonding and hydrophobic interactions could occur. Covalent interactions would not occur. High temperature and high salt concentrations tend to break hydrogen bonds. Therefore, high temperature and high salt would inhibit assembly and stimulate disassembly.

C25. Only the first intron would be spliced out. The mature RNA would be: exon 1–exon 2–intron 2–exon 3.

C26. In bacteria, the 5′ end of the tRNA is cleaved by RNaseP. The 3′ end is cleaved by a different endonuclease, and then a few nucleotides are digested away by an exonuclease that removes nucleotides one at a time until it reaches a CCA sequence.

C27. The spliceosome is composed of multiple protein subunits and some RNAs. The function of the spliceosome is to cut the RNA in two places, hold the ends together, and then catalyze covalent bond formation between the two ends. The role of small snRNAs during this process could be involved in binding the pre-mRNA and/or it could be catalytic, like the RNA in RNaseP.
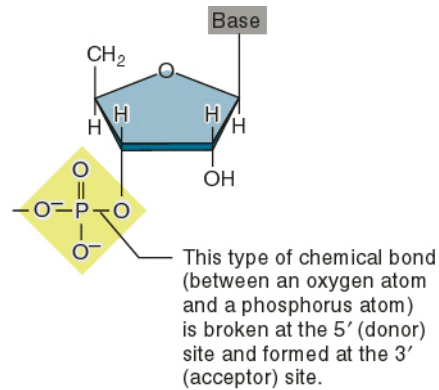
C28. A ribozyme is an enzyme whose catalytic part is composed of RNA. Examples are RNaseP and self-splicing group I and II introns. It is thought that the spliceosome may have catalytic RNAs as well.

C29. A gene is colinear when the sequence of bases in the sense strand of the DNA (i.e., the DNA strand that is complementary to the template strand for RNA synthesis) corresponds to the sequence of bases in the mRNA. Most prokaryotic genes and many eukaryotic genes are colinear. Therefore, you can look at the gene sequence in the DNA and predict the amino acid sequence in the polypeptide. Many eukaryotic genes, however, are not colinear. They contain introns that are spliced out of the pre-mRNA.

C30. Self-splicing means that an RNA molecule can splice itself without the aid of a protein. Group I and II introns can be self-splicing, although proteins can also facilitate the process.

C31. In eukaryotes, pre-mRNA can be capped, tailed, and spliced and then exported out of the nucleus.

C32. A transesterification reaction involves the breakage of one ester bond and formation of another ester bond.

This type of chemical bond (between an oxygen atom and a phosphorus atom) is broken at the 5′ (donor) site and formed at the 3′ (acceptor) site.

C33. Alternative splicing occurs when exons are spliced out or alternative splice sites are used at intron-exon boundaries. The biological significance is that two or more polypeptide sequences can be derived from a single gene. This is a more efficient use of the genetic material. In multicellular organisms, alternative splicing is often used in a cell-specific manner.

C34. This is called trimming or processing rather than splicing because none of the mature rRNA fragments are ever ligated (i.e., connected) to each other after the cleavage of the primary rRNA transcript.

C35. As shown at the *left* side of Figure 12.18, the guanosine, which binds to the guanosine-binding site, does not have a phosphate group attached to it. This guanosine is the nucleoside that winds up at the 5′ end of the intron. Therefore, the intron does not have a phosphate group at its 5′ end.

C36. A. U1 and U4/U6

B. U5

C. U2

C37. U5

C38. The 60-nucleotide sequence would be found in the closed loop. The closed loop is the region between the 5′ splice site and the branch site.