E1. Sticky ends, which are complementary in their DNA sequence, will promote the binding of DNA fragments to each other. This binding is due to hydrogen bonding.

E2. Remember that AT base pairs form two hydrogen bonds while GC base pairs form three hydrogen bonds. The order (from stickiest to least sticky) would be:

$$\textit{Bam}\text{HI} = \textit{Pst}\text{ I} = \textit{Sac}\text{ I} > \textit{Eco}\text{RI} > \textit{Cla}\text{ I}.$$

E3. All vectors have the ability to replicate when introduced into a living cell. This ability is due to a DNA sequence known as an origin of replication. Modern vectors also contain convenient restriction sites for the insertion of DNA fragments. These vectors also contain selectable markers, which are genes that confer some selectable advantage for the host cell that carries them. The most common selectable markers are antibiotic-resistance genes, which confer resistance to antibiotics that would normally inhibit the growth of the host cell.

E4. In conventional gene cloning, many copies are made because the vector replicates to a high copy number within the cell, and the cells divide to produce many more cells. In PCR, the replication of the DNA to produce many copies is facilitated by primers, nucleotides, and *Taq* polymerase.

E5. First, the chromosomal DNA that contains the source of the gene that you want to clone must be obtained from a cell (tissue) sample. A vector must also be obtained. The vector and chromosomal DNA are digested with a restriction enzyme. They are mixed together to allow the sticky ends of the DNA fragments to bind to each other, hopefully to create a hybrid vector. DNA ligase is then added to promote covalent bonds. The DNA is then transformed or transfected into a living cell. The vector will replicate and the cells will divide to produce a colony of cells that contain "cloned" DNA pieces. To identify colonies that contain the gene you wish to clone, you must use a probe that will specifically identify a colony containing the correct hybrid vector. A probe may be a DNA probe that is complementary to the gene you want to clone; or it could be an antibody that recognizes the protein that is encoded by the gene.

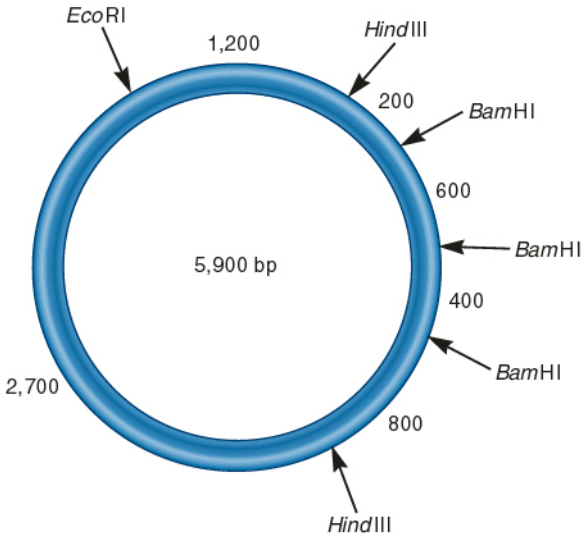
E6. A hybrid vector is a vector that has a piece of "foreign" DNA inserted into it. The foreign DNA came from somewhere else, like the chromosomal DNA of some organism. To construct a hybrid vector, the vector and source of foreign DNA are digested with the same restriction enzyme. The complementary ends of the fragments are allowed to hydrogen bond to each other (i.e., sticky ends are allowed to bind), and then DNA ligase is added to create covalent phosphoester bonds. If all goes well, a piece of the foreign DNA will become ligated to the vector, thereby creating a hybrid vector.

As described in Figure 18.2, the insertion of foreign DNA can be detected using X-Gal. As seen here, the insertion of the foreign DNA causes the inactivation of the *lacZ* gene. The *lacZ* gene encodes the enzyme *β*-galactosidase, which is necessary to convert X-gal to a blue compound. If the *lacZ* gene is inactivated by the insertion of foreign DNA, the bacterial colonies will be white. If the vector has simply recircularized, and the *lacZ* gene remains intact, the colonies will be blue.

E7. Probably not. It was necessary for the researchers to obtain kanamycin-resistant colonies, and this would occur only if the entire kan^R gene were present. If the kan^R gene had been cut in half, it would have been difficult to insert that whole gene into pSC101. It is possible, if the two pieces of pSC101 DNA (carrying the kan^R gene) were inserted into the same pSC101 plasmid. However, the chance of this happening is much less than inserting a single piece of DNA, carrying the entire kan^R gene, into pSC101.

E8. If the *Eco*RI fragment containing the kan^R gene also had an origin of replication, it is possible that this fragment could circularize and become a very small plasmid. The electrophoresis results would be consistent with the idea that the same bacterial cells could contain two different plasmids: pSC101 and a small plasmid corresponding to the segment of DNA (now circularized) that carried the kan^R gene. However, the results shown in step 9 rule out this possibility. The density gradient centrifugation showed a single peak, corresponding to a plasmid that had an intermediate size between pSC101 and pSC102. In contrast, if our alternative explanation had been correct (i.e., that bacterial cells contain two plasmids), there would be two peaks from the density gradient centrifugation. One peak would correspond to pSC101 and the other peak would indicate a very small plasmid (i.e., smaller than pSC101 and pSC102).

E9. The map is shown here.



E10. 3×2^{27}, which equals 4.0×10^{8}, or about 400 million copies.

E11. A thermostable form of DNA polymerase (e.g., *Taq* polymerase) is used in PCR because each PCR cycle involves a heating step to denature the DNA. This heating step would inactivate most forms of DNA polymerase. However, *Taq* polymerase is thermostable and can remain functional after many cycles of heating and cooling. It is not necessary to use a thermostable form of DNA polymerase in the techniques of dideoxy DNA sequence or site-directed mutagenesis. In these methods, DNA polymerase can be added after the annealing step, and the sample can be incubated at a temperature that does not inactivate most forms of DNA polymerase.

E12. Initially, the mRNA would be mixed with reverse transcriptase and nucleotides to create a complementary strand of DNA. Reverse transcriptase also needs a primer. This could be a primer that is known to be complementary to the *ß*-globin mRNA. Alternatively, mature mRNAs have a polyA tail, so one could add a primer that consists of many Ts. This is called a poly-dT primer. After the complementary DNA strand has been made, the sample would then be mixed with primers, *Taq* polymerase, and nucleotides, etc., and subjected to the standard PCR protocol. Note: the PCR reaction would have two kinds of primers. One primer would be complementary to the 5′ end of the mRNA and would be unique to the *ß*-globin sequence. The other primer would be complementary to the 3′ end. This second primer could be a poly-dT primer or it could be a unique primer that would bind slightly upstream from the polyA-tail region.

E13. One interpretation would be that the gene is part of a gene family. In this case, the family would contain four homologous members. At high stringency, the probe binds only to the gene that is its closest match, but at low stringency it recognizes the three other homologous genes.

E14. A DNA library is a collection of hybrid vectors that contain different pieces of DNA from a source of chromosomal DNA. Because it is a diverse collection of many different DNA pieces, the name *library* seems appropriate.

E15. It would be necessary to use cDNA so that the gene would not carry any introns. Bacterial cells do not contain spliceosomes (which are described in Chapter 15). To express a eukaryotic protein in bacteria, a researcher would clone cDNA into bacteria since the cDNA does not contain introns.

E16. Hybridization occurs due to the hydrogen bonding of complementary sequences. Due to the chemical properties of DNA and RNA strands, they form double-stranded regions when the base sequences are complementary. In a Southern and Northern experiment, the probe is labeled.

E17. The purpose of gel electrophoresis is to separate the many DNA fragments, RNA molecules, or proteins that were obtained from the sample you want to probe. This separation is based on molecular mass and allows you to identify the molecular mass of the DNA fragment, RNA molecule, or protein that is being recognized by the probe.

E18. The purpose of a Northern blot experiment is to determine if a gene is transcribed into RNA using a piece of cloned DNA as a probe. It can tell you if a gene is transcribed in a particular cell or at a particular stage of development. It can also tell you if a pre-mRNA is alternatively spliced.

E19. The Northern blot is shown here. The female mouse expresses the same total amount of this mRNA compared to the male. In the heterozygous female, there would be 50% of the 900 bp band and 50% of the 825 bp band.

E20. It appears that this mRNA is alternatively spliced to create a high molecular mass and a lower molecular mass product. Nerve cells produce a very large amount of the larger mRNA, whereas spleen cells produce a moderate amount of the smaller mRNA. Both types are produced in small amounts by the muscle cells. It appears that kidney cells do not transcribe this gene.

E21. Restriction enzymes recognize many sequences throughout the chromosomal DNA. If two fragments from different samples have the same molecular mass in a Southern blot, it is likely (though not certain) that the two fragments are found at the same chromosomal site in the genome. In this Southern blot, most of the transposable elements are found at the same sites within the genomes of these different yeast strains. However, a couple of bands are different among the three strains. These results indicate that the *Ty* element may occasionally transpose to a new location or that chromosomal changes (point mutations, chromosomal rearrangements, deletions, etc.) may have slightly changed the genomes among these three strains of yeast in a way that changes the distances between the restriction sites.

E22. 1. *β* (detected at the highest stringency)

2. δ
3. γ_A and ε
4. α_1
5. *Mb* (detected only at the lowest stringency)

Note: At the lowest stringency, all of the globin genes would be detected. At the highest stringency, only the *β*-globin gene would be detected.

E23. Lane 1 shows that *β*-globin is made in normal red blood cells. In red blood cells from a thalassemia patient (lane 2), however, very little is made. Perhaps this person is homozygous for a down promoter mutation, which diminishes the transcription of the gene. As shown in lanes 3 and 4, *β*-globin is not made in muscle cells.

E24. The Western blot is shown here. The sample in lane 2 came from a plant that was homozygous for a mutation that prevented the expression of this polypeptide. Therefore, no protein was observed in this lane. The sample in lane 4 came from a plant that is homozygous for a mutation that introduces an early stop codon into the coding sequence. As seen in lane 4, the polypeptide is shorter than normal (13.3 kDa). The sample in lane 3 was from a heterozygote that expresses about 50% of each type of polypeptide. Finally, the sample in lane 5 came from a plant that is homozygous for a mutation that changed one amino acid to another amino acid. This type of mutation, termed a missense mutation, may not be detectable on gel. However, a single amino acid substitution could affect polypeptide function.

E25. Western blotting.

E26. The products of structural genes are proteins with a particular amino acid sequence. Antibodies can specifically recognize proteins due to their amino acid sequence. Therefore, an antibody can detect whether or not a cell is making a particular type of protein.

E27. You would first make a DNA library using chromosomal DNA from pig cells. You would then radiolabel the human *ß*-globin gene and use it as a probe in a colony hybridization experiment; each colony would contain a different cloned piece of the pig genome. You would identify "hot" colonies that hybridize to the human *ß*-globin probe. You would then go back to the master plate and pick these colonies and grow them in a test tube. You would then isolate the plasmid DNA from the colonies and subject the DNA to DNA sequencing. By comparing the DNA sequences of the human *ß*-globin gene and the putative clones, you could determine if the putative clones were homologous to the human clone and likely to be the pig homologue of the *ß*-globin gene.

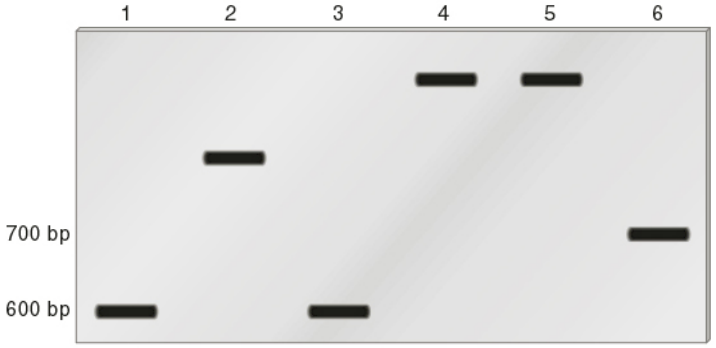
E28. In this case, the transcription factor binds to the response element when the hormone is present. Therefore, the hormone promotes the binding of the transcription factor to the DNA and thereby promotes transactivation.

E29. The rationale behind a gel retardation assay is that a segment of DNA with a protein bound to it will migrate more slowly through a gel than will the same DNA without any bound protein. A shift in a DNA band to a higher molecular mass provides a way to identify DNA-binding proteins.

E30. The levels of cAMP affect the phosphorylation of CREB, and this affects whether or not it can transactivate transcription. However, CREB can bind to CREs whether or not it is phosphorylated. Therefore, in a gel retardation assay, we would expect CREB to bind to CREs and retard their mobility using a cell extract from cells that were or were not retreated with adrenalin.

E31. TFIID can bind to this DNA fragment by itself, as seen in lane 2. However, TFIIB and RNA polymerase II cannot bind to the DNA by themselves (lanes 3 and 4). As seen in lane 5, TFIIB can bind, if TFIID is also present, because the mobility shift is higher than TFIID alone (compare lanes 2 and 5). In contrast, RNA polymerase II cannot bind to the DNA when only TFIID is present. The mobility shift in lane 6 is the same as that found in lane 2, indicating that only TFIID is bound. Finally, in lane 7, when all three components are present, the mobility shift is higher than when TFIIB and TFIID are present (compare lanes 5 and 7). These results mean that all three proteins are bound to the DNA. Taken together, the results indicate that TFIID can bind by itself, TFIIB needs TFIID to bind, and RNA polymerase II needs both proteins to bind to the DNA.

E32. The glucocorticoid receptor will bind to GREs if glucocorticoid hormone is also present. The glucocorticoid receptor does not bind to CREs. The CREB protein will bind to CREs (with or without hormone), but it will not bind to GREs. The expected results are shown here. In this drawing, the binding of CREB protein to the 700 bp fragment results in a complex with a higher mass compared to the glucocorticoid receptor binding to the 600 bp fragment.



E33. The region of the gel from about 350 bp to 175 bp does not contain any bands. This is the region being covered up; it is about 175 bp long.

E34. The rationale behind a footprinting experiment has to do with accessibility. If a protein is bound to the DNA, it will cover up the part of the DNA where it is bound. This region of the DNA will be inaccessible to the actions of chemicals or enzymes that cleave the DNA, such as DNase I.

E35. The A closest to the bottom of the gel is changed to a G.

E36. A. AGGTCGGTTGCCATCGCAATAATTTCTGCCTGAACCCAATA

B. Automated sequencing has several advantages. First, the reactions are done in a single tube as opposed to four tubes. Second, the detector can "read" the sequence and provide the researcher with a printout of the sequence. This is much easier than looking at an X-ray film and writing the sequence out by hand. It also avoids human error. Finally, automated sequencing does not require the use of radioisotopes, which are more expensive and require more laboratory precautions, compared to fluorescently labeled compounds.

E37. 5′–CCCCCGATC**G**ACATCATTA–3′. The mutagenic base is underlined.

E38. There are lots of different strategies one could follow. For example, you could mutate every other base and see what happens. It would be best to make very nonconservative mutations such as a purine for a pyrimidine or a pyrimidine for a purine. If the mutation prevents protein binding in a gel retardation assay, then the mutation is probably within the response element. If the mutation has no effect on protein binding, it probably is outside the response element.

E39. You would conclude that it might be important. The only amino acid substitution that gave a substantial amount of functional activity was an aspartate. Glutamate and aspartate have very similar amino acid side chains (see Chapter 13); they both contain a carboxyl (COOH) group. Based on these results, you would suspect that a carboxyl group at this location in the protein might be important for its function.