CHAPTER

6 The Structures of DNA and RNA

The discovery that DNA is the prime genetic molecule, carrying all the hereditary information within chromosomes, immediately focused attention on its structure. It was hoped that knowledge of the structure would reveal how DNA carries the genetic messages that are replicated when chromosomes divide to produce two identical copies of themselves. During the late 1940s and early 1950s, several research groups in the United States and in Europe engaged in serious efforts—both cooperative and rival—to understand how the atoms of DNA are linked together by covalent bonds and how the resulting molecules are arranged in three-dimensional space. Not surprisingly, there initially were fears that DNA might have very complicated and perhaps bizarre structures that differed radically from one gene to another. Great relief, if not general elation, was thus expressed when the fundamental DNA structure was found to be the double helix. It told us that all genes have roughly the same three-dimensional form and that the differences between two genes reside in the order and number of their four nucleotide building blocks along the complementary strands.

Now, some 50 years after the discovery of the double helix, this simple description of the genetic material remains true and has not had to be appreciably altered to accommodate new findings. Nevertheless, we have come to realize that the structure of DNA is not quite as uniform as was first thought. For example, the chromosome of some small viruses have single-stranded, not double-stranded, molecules. Moreover, the precise orientation of the base pairs varies slightly from base pair to base pair in a manner that is influenced by the local DNA sequence. Some DNA sequences even permit the double helix to twist in the left-handed sense, as opposed to the right-handed sense originally formulated for DNA's general structure. And while some DNA molecules are linear, others are circular. Still additional complexity comes from the supercoiling (further twisting) of the double helix, often around cores of DNA-binding proteins.

Likewise, we now realize that RNA, which at first glance appears to be very similar to DNA, has its own distinctive structural features. It is principally found as a single-stranded molecule. Yet by means of intra-strand base pairing, RNA exhibits extensive double-helical character and is capable of folding into a wealth of diverse tertiary structures. These structures are full of surprises, such as non-classical base pairs, base-backbone interactions, and knot-like configurations. Most remarkable of all, and of profound evolutionary significance, some RNA molecules are enzymes that carry out reactions that are at the core of information transfer from nucleic acid to protein.

Clearly, the structures of DNA and RNA are richer and more intricate than was at first appreciated. Indeed, there is no one generic structure for DNA and RNA. As we shall see in this chapter, there are in fact variations on common themes of structure that arise from the unique physical, chemical, and topological properties of the polynucleotide chain. ■

OUTLINE

- DNA Structure (p. 2)
- DNA Topology (p. 17)
- RNA Structure (p. 25)

DNA STRUCTURE

DNA Is Composed of Polynucleotide Chains

The most important feature of DNA is that it is usually composed of two **polynucleotide chains** twisted around each other in the form of a double helix (Figure 6-1). The upper part of the figure (a) presents the structure of the double helix shown in a schematic form. Note that if inverted 180° (for example, by turning this book upside-down), the double helix looks superficially the same, due to the complementary nature of the two DNA strands. The space-filling model of the double helix, in the lower part of the figure (b), shows the components of the DNA molecule and their relative positions in the helical structure. The backbone of each strand of the helix is composed of alternating sugar and phosphate residues; the bases project inward but are accessible through the major and minor grooves.

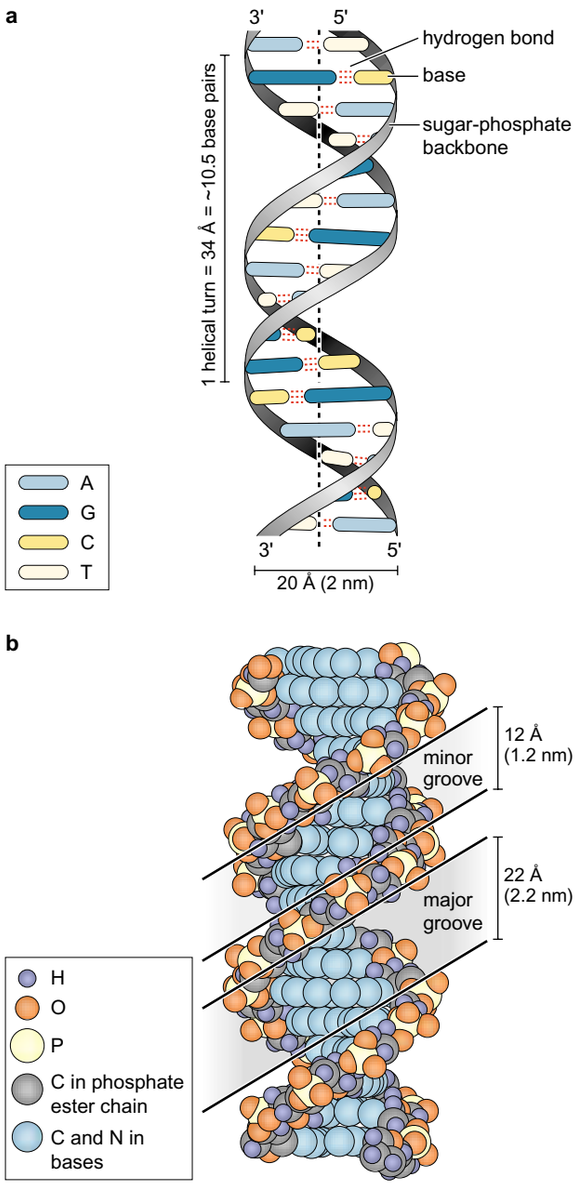


FIGURE 6-1 The Helical Structure of DNA. (a) Schematic model of the double helix. One turn of the helix (34 Å or 3.4 nm) spans approx. 10.5 base pairs. (b) Space-filling model of the double helix. The sugar and phosphate residues in each strand form the backbone, which are traced by the yellow, gray, and red circles, show the helical twist of the overall molecule. The bases project inward but are accessible through major and minor grooves.

Let us begin by considering the nature of the nucleotide, the fundamental building block of DNA. The nucleotide consists of a phosphate joined to a sugar, known as **2′-deoxyribose,** to which a base is attached. The phosphate and the sugar have the structures shown in Figure 6-2. The sugar is called *2′*-deoxyribose because there is no hydroxyl at position 2′ (just two hydrogens). Note that the positions on the ribose are designated with primes to distinguish them from positions on the bases (see the discussion below).

We can think of how the base is joined to 2′-deoxyribose by imagining the removal of a molecule of water between the hydroxyl on the 1′ carbon of the sugar and the base to form a glycosidic bond (Figure 6-2). The sugar and base alone are called a **nucleoside.** Likewise, we can imagine linking the phosphate to 2′-deoxyribose by removing a water molecule from between the phosphate and the hydroxyl on the 5′ carbon to make a 5′ phosphomonoester. Adding a phosphate (or more than one phosphate) to a **nucleoside** creates a **nucleotide.** Thus, by making a glycosidic bond between the base and the sugar, and by making a phosphoester bond between the sugar and the phosphoric acid, we have created a nucleotide (Table 6-1).

Nucleotides are, in turn, joined to each other in polynucleotide chains through the 3′ hydroxyl of 2′-deoxyribose of one nucleotide and the phosphate attached to the 5′ hydroxyl of another nucleotide (Figure 6-3). This is a **phosphodiester linkage** in which the phosphoryl group between the two nucleotides has one sugar esterified to it through a 3′ hydroxyl and a second sugar esterified to it through a 5′ hydroxyl. Phosphodiester linkages create the repeating, sugar-phosphate backbone of the polynucleotide chain, which is a regular feature of DNA. In contrast, the order of the bases along the polynucleotide chain is irregular. This irregularity as well as the long length is, as we shall see, the basis for the enormous information content of DNA.

The phosphodiester linkages impart an inherent polarity to the DNA chain. This polarity is defined by the asymmetry of the nucleotides and the way they are joined. DNA chains have a free 5′ phosphate or 5′ hydroxyl at one end and a free 3′ phosphate or 3′ hydroxyl at the other end. The convention is to write DNA sequences from the 5′ end (on the left) to the 3′ end, generally with a 5′ phosphate and a 3′ hydroxyl.

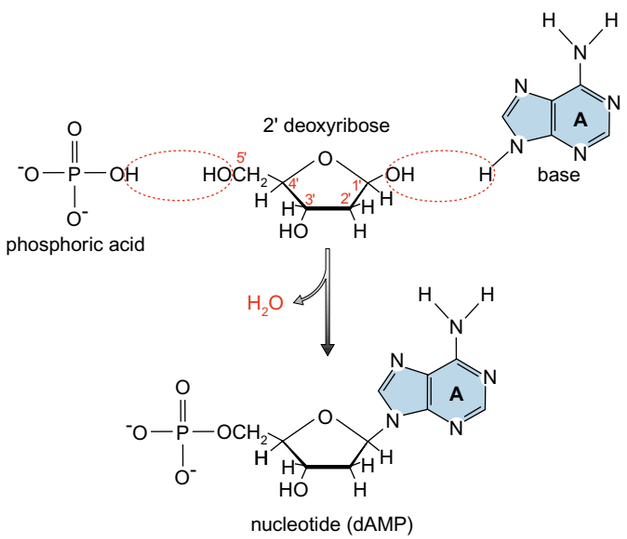
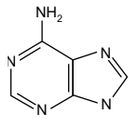
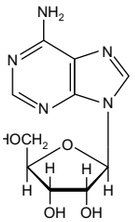
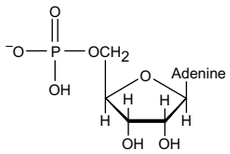
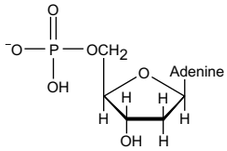


FIGURE 6-2 Formation of Nucleotide by Removal of Water. The numbers of the carbon atoms in 2' deoxyribose are labeled in red.

TABLE 6-1 Adenine and Related Compounds

	Base Adenine	Nucleoside Adenosine	Nucleotide Adenosine 5′-phosphate	Deoxynucleotide Deoxyadenosine 5′ phosphate
Structure[a]				
M.W.	135.1	267.2	347.2	331.2

[a]At physiological pH, all of the hydroxyls bound to phosphate are ionized.

Each Base Has Its Preferred Tautomeric Form

The bases in DNA fall into two classes, **purines** and **pyrimidines.** The purines are **adenine** and **guanine,** and the pyrimidines are **cytosine** and **thymine.** The purines are derived from the double-ringed structure shown in Figure 6-4. Adenine and guanine share this essential structure but with different groups attached. Likewise, cytosine and thymine are

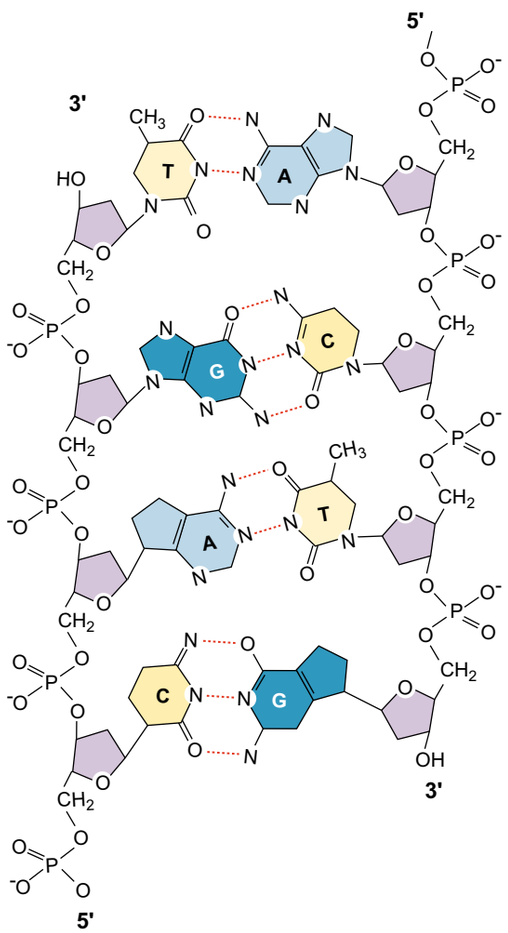


FIGURE 6-3 Detailed Structure of Polynucleotide Polymer. The structure shows base pairing between purines (in blue) and pyrimidines (in yellow), and the phosphodiester linkages of the backbone.

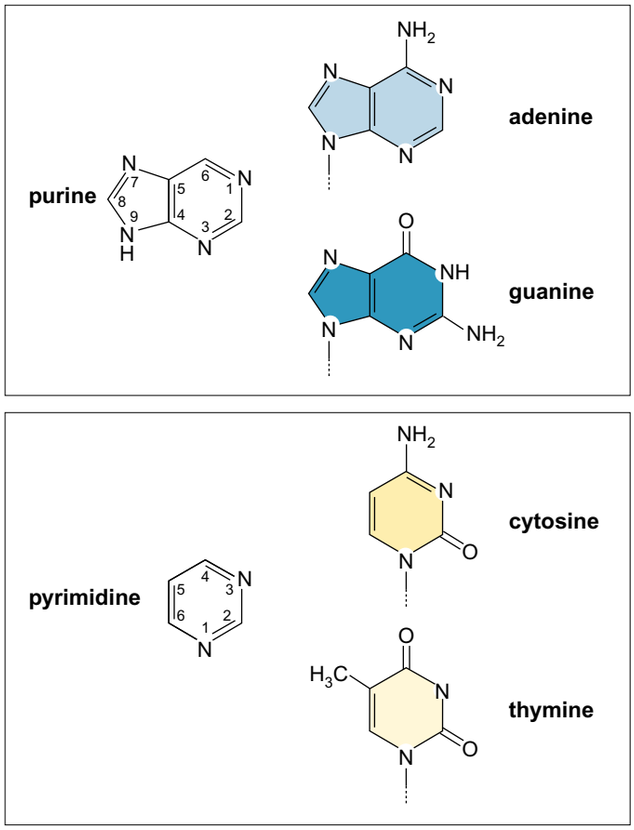


FIGURE 6-4 Purines and Pyrimidines. The dotted lines indicate the sites of attachment of the bases to the sugars.

variations on the single-ringed structure shown in Figure 6-4. The figure also shows the numbering of the positions in the purine and pyrimidine rings. The bases are attached to the deoxyribose by glycosidic linkages at N1 of the pyrimidines or at N9 of the purines.

Each of the bases exists in two alternative **tautomeric states,** which are in equilibrium with each other. The equilibrium lies far to the side of the conventional structures shown in Figure 6-4, which are the predominant states and the ones important for base pairing. The nitrogen atoms attached to the purine and pyrimidine rings are in the amino form in the predominant state and only rarely assume the imino configuration. Likewise, the oxygen atoms attached to the guanine and thymine normally have the keto form and only rarely take on the enol configuration. As examples, Figure 6-5 shows tautomerization of cytosine into the imino form (a) and guanine into the enol form (b). As we shall see, the capacity to form an alternative tautomer is a frequent source of errors during DNA synthesis.

The Two Strands of the Double Helix Are Held Together by Base Pairing in an Anti-Parallel Orientation

The double helix is composed of two polynucleotide chains that are held together by weak, non-covalent bonds between pairs of bases, as shown in Figure 6-3. Adenine on one chain is always paired with thymine on the other chain and, likewise, guanine is always paired with cytosine. The two strands have the same helical geometry but base pairing holds them together with the opposite polarity. That is, the base at the 5′ end of one strand is paired with the base at the 3′ end of the other strand. The strands are said to have an anti-parallel

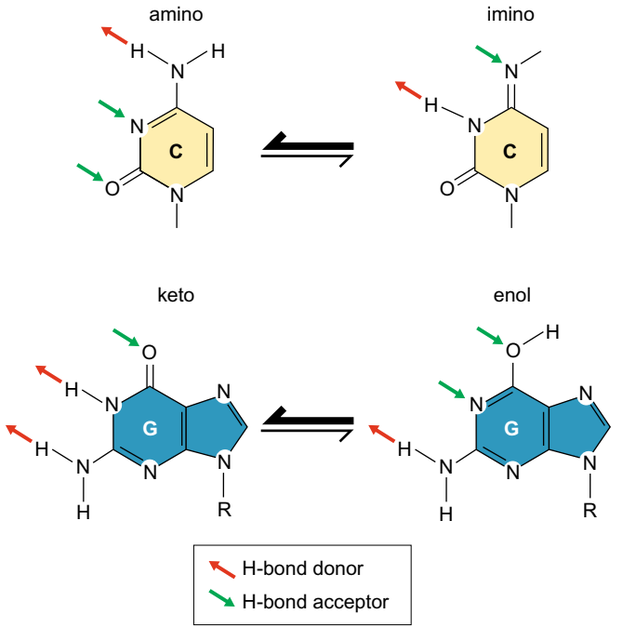


FIGURE 6-5 Base Tautomers.
Amino ⇋ imino and keto ⇋ enol tautomerism. (a) Cytosine is usually in the amino form but rarely forms the imino configuration. (b) Guanine is usually in the keto form but is rarely found in the enol configuration.

orientation. This anti-parallel orientation is a stereochemical consequence of the way that adenine and thymine and guanine and cytosine pair with each together (see Figure 6-6).

The Two Chains of the Double Helix Have Complementary Sequences

The pairing between adenine and thymine and between guanine and cytosine results in a complementary relationship between the sequence of bases on the two intertwined chains and gives DNA its self-encoding character. For example, if we have the sequence 5′-ATGTC-3′ on one chain, the opposite chain must have the complementary sequence 3′-TACAG-5′.

The strictness of the rules for this "Watson-Crick" pairing derives from the complementarity both of shape and of hydrogen bonding properties between adenine and thymine and between guanine and cytosine (Figure 6-6). Adenine and thymine match up so that a hydrogen bond can form between the exocyclic amino group at C6 on adenine and the carbonyl at C4 in thymine; and likewise, a hydrogen bond can form between N1 of adenine and N3 of thymine. A corresponding arrangement can be drawn between a guanine and a cytosine, so that there is both hydrogen bonding and shape complementarity in this base pair as well. A G:C base pair has three hydrogen bonds, because the exocyclic NH_2 at C2 on guanine lies opposite to, and can hydrogen bond with, a carbonyl at C2 on cytosine. Likewise, a hydrogen bond can form between N1 of guanine and N3 of cytosine and between the carbonyl at C6 of guanine and the exocyclic NH_2 at C4 of cytosine. Watson-Crick base pairing requires that the bases are in their preferred tautomeric states.

An important feature of the double helix is that the two base pairs have exactly the same geometry; having an A:T base pair or a G:C base pair between the two sugars does not perturb the arrangement of the sugars. Neither does T:A or C:G. In other words, there is an approximately twofold axis of symmetry that relates the two sugars and all

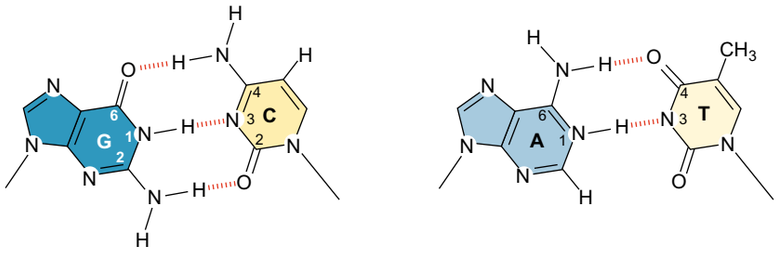


FIGURE 6-6 A:T and G:C Base Pairs. The figure shows hydrogen bonding between the bases.

four base pairs can be accommodated within the same arrangement without any distortion of the overall structure of the DNA.

Hydrogen Bonding Is Important for the Specificity of Base Pairing

The hydrogen bonds between complementary bases are a fundamental feature of the double helix, contributing to the thermodynamic stability of the helix and providing the information content and specificity of base pairing. Hydrogen bonding might not at first glance appear to contribute importantly to the stability of DNA for the following reason. An organic molecule in aqueous solution has all of its hydrogen bonding properties satisfied by water molecules that come on and off very rapidly. As a result, for every hydrogen bond that is made when a base pair forms, a hydrogen bond with water is broken that was there before the base pair formed. Thus, the net energetic contribution of hydrogen bonds to the stability of the double helix would appear to be modest. However, when polynucleotide strands are separate, water molecules are lined up on the bases. When strands come together in the double helix, the water molecules are displaced from the bases. This creates disorder and increases entropy, thereby stabilizing the double helix. Hydrogen bonds are not the only force that stabilizes the double helix. A second important contribution comes from stacking interactions between the bases. The bases are flat, relatively water-insoluble molecules, and they tend to stack above each other roughly perpendicular to the direction of the helical axis. Electron cloud interactions ($\pi-\pi$) between bases in the helical stacks contribute significantly to the stability of the double helix.

Hydrogen bonding is also important for the specificity of base pairing. Suppose we tried to pair an adenine with a cytosine. Then we would have a hydrogen bond acceptor (N1 of adenine) lying opposite a hydrogen bond acceptor (N3 of cytosine) with no room to put a water molecule in between to satisfy the two acceptors (Figure 6-7). Likewise, two hydrogen bond donors, the NH_2 groups at C6 of adenine and C4 of cytosine, would lie opposite each other. Thus, an A:C base pair would be unstable because water would have to be stripped off the donor and acceptor groups without restoring the hydrogen bond formed within the base pair.

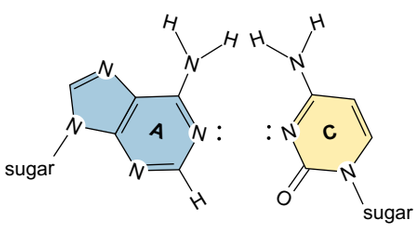


FIGURE 6-7 A:C Incompatibility. The structure shows the inability of adenine to form the proper hydrogen bonds with cytosine. The base pair is therefore unstable.

FIGURE 6-8 Base Flipping. Structure of isolated DNA, showing the flipped cytosine residue and the small distortions to the adjacent base pairs. (Source: Reprinted/redrawn from Roberts, R. J. 1995. *Cell* **82**(1):9–12.)

Bases Can Flip Out from the Double Helix

As we have seen, the energetics of the double helix favor the pairing of each base on one polynucleotide strand with the complementary base on the other strand. Sometimes, however, individual bases can protrude from the double helix in a remarkable phenomenon known as **base flipping** shown in Figure 6-8. As we shall see in Chapter 9, certain enzymes that methylate bases or remove damaged bases do so with the base in an extra helical configuration in which it is flipped out from the double helix, enabling the base to sit in the catalytic cavity of the enzyme. Furthermore, enzymes involved in homologous recombination and DNA repair are believed to scan DNA for homology or lesions by flipping out one base after another. This is not energetically expensive because only one base is flipped out at a time. Clearly, DNA is more flexible than might be assumed at first glance.

DNA Is Usually a Right-Handed Double Helix

Applying the handedness rule from physics, we can see that each of the polynucleotide chains in the double helix is right-handed. In your mind's eye, hold your right hand up to the DNA molecule in Figure 6-9 with your thumb pointing up and along the long axis of the helix and your fingers following the grooves in the helix. Trace along one strand of the helix in the direction in which your thumb is pointing. Notice that you go around the helix in the same direction as your fingers are pointing. This does not work if you use your left hand. Try it!

A consequence of the helical nature of DNA is its periodicity. Each base pair is displaced (twisted) from the previous one by about 36°. Thus, in the X-ray crystal structure of DNA it takes a stack of about 10 base pairs to go completely around the helix (360°) (see Figure 6-1a). That is, the helical periodicity is generally 10 base pairs per turn of the helix. For further discussion, see Box 6-1: DNA Has 10.5 Base Pairs per Turn of the Helix in Solution: The Mica Experiment.

The Double Helix Has Minor and Major Grooves

As a result of the double-helical structure of the two chains, the DNA molecule is a long extended polymer with two grooves that are not equal in size to each other. Why are there a minor groove and a major groove? It is a simple consequence of the geometry of the base pair. The angle at which the two sugars protrude from the base pairs (that is, the angle between the glycosidic bonds) is about 120° (for the narrow angle or 240° for the wide angle) (see Figures 6-1b and 6-6). As a result, as more and more base pairs stack on top of each other, the

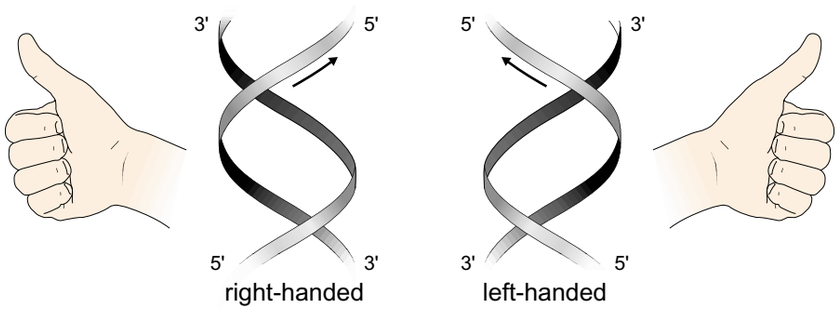
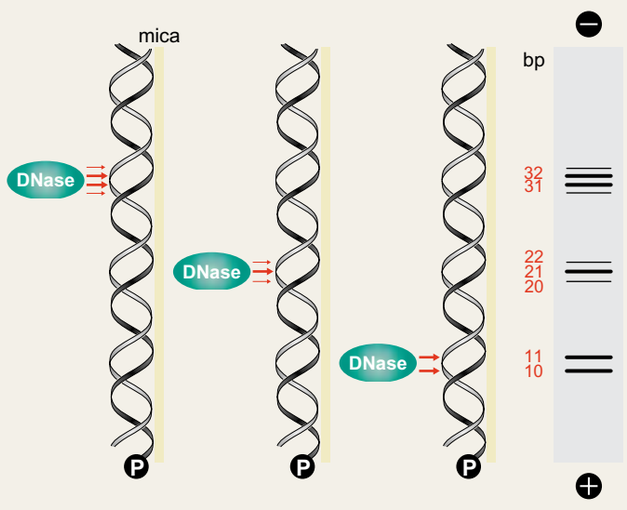


FIGURE 6-9 Left- and Right-Handed Helices. Please see text for details.

Box 6-1 DNA Has 10.5 Base Pairs per Turn of the Helix in Solution: The Mica Experiment

This value of 10 base pairs per turn varies somewhat under different conditions. A classic experiment that was carried out in the 1970s demonstrated that DNA adsorbed on a surface has somewhat greater than 10 base pairs per turn. Short segments of DNA were allowed to bind to mica surface. The presence of 5′-terminal phosphates on the DNAs held them in a fixed orientation on the mica. The mica-bound DNAs were then exposed to DNase I, an enzyme (a deoxyribonuclease) that cleaves the phosphodiester bonds in the DNA backbone. Because the enzyme is bulky, it is only able to cleave phosphodiester bonds on the DNA surface furthest from the mica (think of the DNA as a cylinder lying down on a flat surface) due to the steric difficulty of reaching the sides or bottom surface of the DNA. As a result, the length of the resulting fragments should reflect the periodicity of the DNA, the number of base pairs per turn.

After the mica-bound DNA was exposed to DNase the resulting fragments were separated by electrophoresis in a polyacrylamide gel, a jelly-like matrix (Box 6-1 Figure 1). Because DNA is negatively charged, it migrates through the gel toward the positive pole of the electric field. The gel matrix impedes movement of the fragments in a manner that is proportional to their length such that larger fragments migrate more slowly than smaller fragments. When the experiment is carried out, we see clusters of DNA fragments of average sizes 10 and 11, 21, 31 and 32 base pairs and so forth, that is, in multiples of 10.5, which is the number of base pairs per turn. This value of 10.5 base pairs per turn is close to that of DNA in solution as inferred by other methods (see the section titled The Double Helix Exists in Multiple Conformations, below). The strategy of using DNase to probe the structure of DNA is now used to analyze the interaction of DNA with proteins (see Chapter 17).



BOX 6-1 FIGURE 1 The Mica Experiment.

narrow angle between the sugars on one edge of the base pairs generates a **minor groove** and the large angle on the other edge generates a **major groove.** (If the sugars pointed away from each other in a straight line, that is, at an angle of 180°, then two grooves would be of equal dimensions and there would be no minor and major grooves.)

The Major Groove is Rich in Chemical Information

The edges of each base pair are exposed in the major and minor grooves, creating a pattern of hydrogen bond donors and acceptors and of van der Waals surfaces that identifies the base pair (see Figure 6-10). The edge of an A:T base pair displays the following chemical groups in the following order in the major groove: a hydrogen bond acceptor (the N7 of adenine), a hydrogen bond donor (the exocyclic amino group on C6 of adenine), a hydrogen bond acceptor (the carbonyl group on C4 of thymine) and a bulky hydrophobic surface (the methyl group on C5 of thymine). Similarly, the edge of a G:C base pair displays the following groups in the major groove: a hydrogen bond acceptor (at N7 of guanine), a hydrogen bond acceptor (the carbonyl on C6 of guanine), a hydrogen bond donor (the exocyclic amino group on C4 of cytosine), a small non-polar hydrogen (the hydrogen at C5 of cytosine).

Thus, there are characteristic patterns of hydrogen bonding and of overall shape that are exposed in the major groove that distinguish an A:T base pair from a G:C base pair, and, for that matter, A:T from T:A, and G:C from C:G. We can think of these features as a code in which **A** represents a **hydrogen bond acceptor, D** a **hydrogen bond donor, M** a **methyl group,** and **H** a **nonpolar hydrogen.** In such a code, **A D A M**

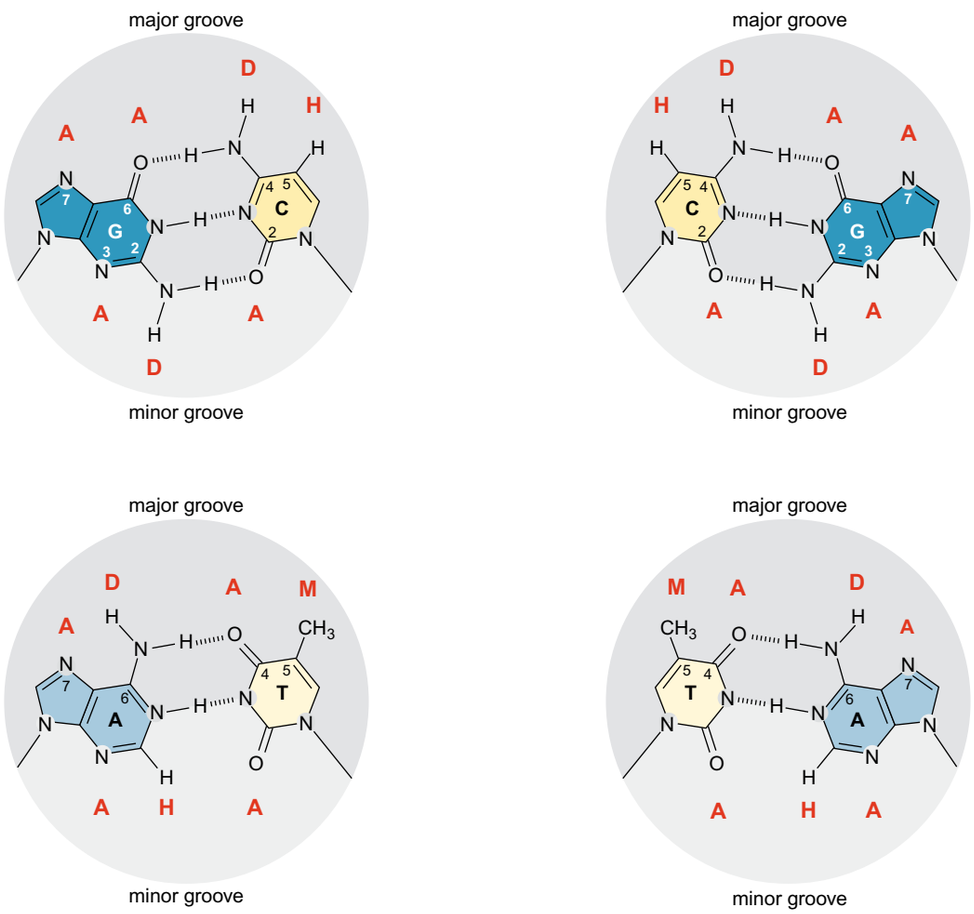


FIGURE 6-10 Chemical Groups Exposed in the Major and Minor Grooves from the Edges of the Base Pairs. The letters in red identify hydrogen bond acceptors (A), hydrogen bond donors (D), nonpolar hydrogens (H), and methyl groups (M).

in the major groove signifies an A:T base pair, and **A A D H** stands for a G:C base pair. Likewise, **M A D A** stands for a T:A base pair and **H D A A** is characteristic of a C:G base pair. In all cases, this code of chemical groups in the major groove specifies the identity of the base pair. These patterns are important because they allow proteins to unambiguously recognize DNA sequences without having to open and thereby disrupt the double helix. Indeed, as we shall see, a principal decoding mechanism relies upon the ability of amino acid side chains to protrude into the major groove and to recognize and bind to specific DNA sequences.

The minor groove is not as rich in chemical information and what information is available is less useful for distinguishing between base pairs. The small size of the minor groove is less able to accommodate amino acid side chains. Also, A:T and T:A base pairs and G:C and C:G pairs look similar to one another in the minor groove. An A:T base pair has a hydrogen bond acceptor (at N3 of adenine), a nonpolar hydrogen (at N2 of adenine) and a hydrogen bond acceptor (the carbonyl on C2 of thymine). Thus, its code is **A H A**. But this code is the same if read in the opposite direction, and hence an A:T base pair does not look very different from a T:A base pair from the point of view of the hydrogen-bonding properties of a protein poking its side chains into the minor groove. Likewise, a G:C base pair exhibits a hydrogen bond acceptor (at N3 of guanine), a hydrogen bond donor (the exocyclic amino group on C2 of guanine), and a hydrogen bond acceptor (the carbonyl on C2 of cytosine), representing the code **A D A**. Thus, from the point of view of hydrogen bonding, C:G and G:C base pairs do not look very different from each other either. The minor groove does look different when comparing an A:T base pair with a G:C base pair, but G:C and C:G, or A:T and T:A, cannot be easily distinguished (see Figure 6-10).

The Double Helix Exists in Multiple Conformations

Early X-ray diffraction studies of DNA, which were carried out using concentrated solutions of DNA that had been drawn out into thin fibers, revealed two kinds of structures, the B and the A forms of DNA (Figure 6-11). The B form, which is observed at high humidity, most closely corresponds to the average structure of DNA under physiological conditions. It has 10 base pairs per turn, and a wide major groove and a narrow minor groove. The A form, which is observed under conditions of low humidity, has 11 base pairs per turn. Its major groove is narrower and much deeper than that of the B form, and its minor groove is broader and shallower. The vast majority of the DNA in the cell is in the B form, but DNA does adopt the A structure in certain DNA-protein complexes. Also, as we shall see, the A form is similar to the structure that RNA adopts when double helical.

The B form of DNA represents an ideal structure that deviates in two respects from the DNA in cells. First, DNA in solution, as we have seen, is somewhat more twisted on average than the B form, having on average 10.5 base pairs per turn of the helix. Second, the B form is an average structure whereas real DNA is not perfectly regular. Rather, it exhibits variations in its precise structure from base pair to base pair. This was revealed by comparison of the crystal structures of individual DNAs of different sequences. For example, the two members of each base pair do not always lie exactly in the same plane. Rather, they can display a “propeller twist” arrangement in which the two flat bases counter rotate relative to each other along the long axis of the base pair,

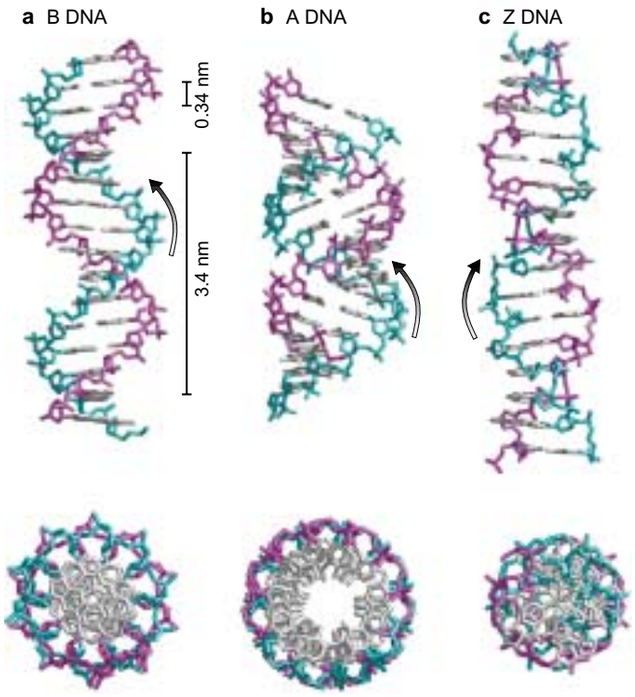


FIGURE 6-11 Models of the B, A, and Z Forms of DNA. The sugar-phosphate backbone of each chain is on the outside in all structures (one red and one blue) with the bases (silver) oriented inward. Side views are shown at the top, and views along the helical axis at the bottom. (a) The B form of DNA, the usual form found in cells, is characterized by a helical turn every 10 base pairs (3.4 nm); adjacent stacked base pairs are 0.34 nm apart. The major and minor grooves are also visible. (b) The more compact A form of DNA has 11 base pairs per turn and exhibits a large tilt of the base pairs with respect to the helix axis. In addition, the A form has a central hole (bottom). This helical form is adopted by RNA–DNA and RNA–RNA helices. (c) Z DNA is a left-handed helix and has a zig zag (hence "Z") appearance. [Courtesy of C. Kielkopf and P. B. Dervan.]

giving the base pair a propeller-like character (Figure 6-12). Moreover, the precise rotation per base pair is not a constant. As a result, the width of the major and minor grooves varies locally. Thus, DNA molecules are never perfectly regular double helices. Instead, their exact conformation depends on which base pair (A:T, T:A, G:C, or C:G) is present at each position along the double helix and on the identity of neighboring base pairs. Still, the B form is for many purposes a good first approximation of the structure of DNA in cells.

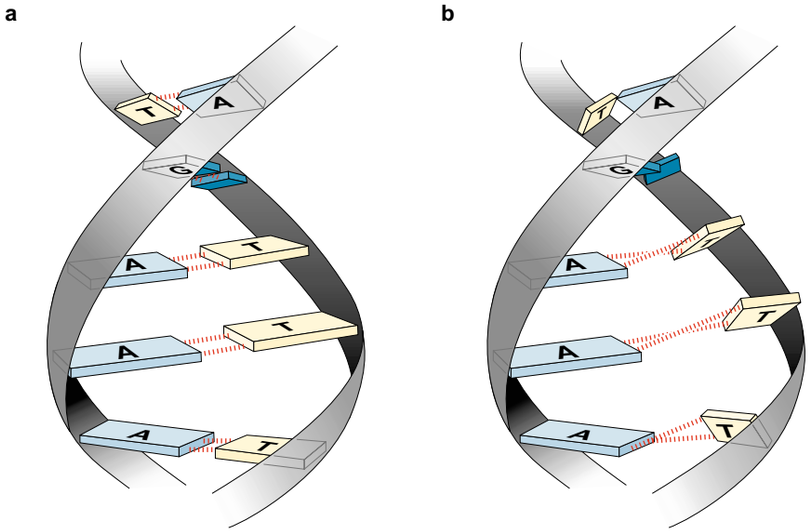


FIGURE 6-12 The Propeller Twist between the Purine and Pyrimidine Base Pairs of a Right-Handed Helix. (a) The structure shows a sequence of three consecutive A:T base pairs with normal Watson-Crick bonding. (b) Propeller twist causes rotation of the bases about their long axis. (Source: Adapted from Aggaarwal et al. 1988. *Science* **242**:899–907, Figure 5B.)

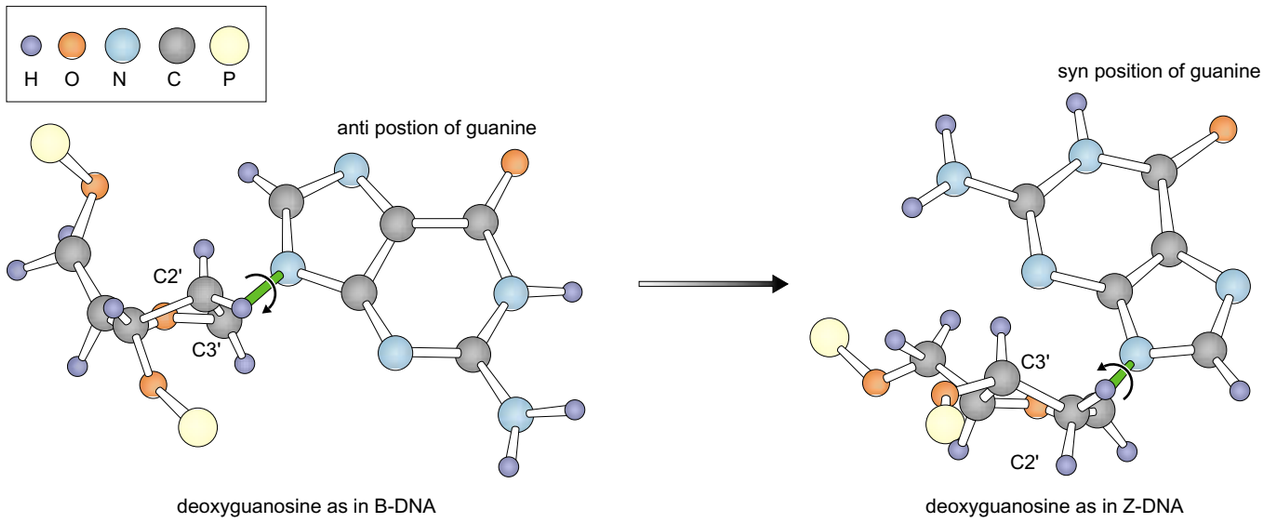


FIGURE 6-13 ***Syn* and *Anti* Positions of Guanine in B and Z DNA.** In right-handed B DNA, the glycosyl bond (colored green) connecting the base to the deoxyribose group is always in the *anti* position, while in left-handed Z DNA it rotates in the direction of the arrow, forming the *syn* conformation at the purine (here guanine) residues but remains in the regular *anti* position (no rotation) in the pyrimidine residues. (Source: After Wang, A. H. J. et al.1983. *Cold Spring Harbor Symp. Quant. Biol* **47**, p. 41.)

DNA Can Sometimes Form a Left-Handed Helix

DNA containing alternative purine and pyrimidine residues can fold into left-handed as well as right-handed helices. To understand how DNA can form a left-handed helix, we need to consider the glycosidic bond that connects the base to the 1′ position of 2′-deoxyribose. This bond can be in one of two conformations called ***syn*** and ***anti*** (Figure 6-13). In right-handed DNA, the glycosidic bond is always in the *anti* conformation. In the left-handed helix, the fundamental repeating unit usually is a purine-pyrimidine dinucleotide, with the glycosidic bond in the *anti* conformation at pyrimidine residues and in the *syn* conformation at purine residues. It is this *syn* conformation at the purine nucleotides that is responsible for the left-handedness of the helix. The change to the *syn* position in the purine residues to alternating *anti–syn* conformations gives the backbone of left-handed DNA a zigzag look (hence its designation of **Z DNA;** see Figure 6-11), which distinguishes it from right-handed forms. The rotation that effects the change from *anti* to *syn* also causes the ribose group to undergo a change in its pucker. Note, as shown in Figure 6-13, that C3′ and C2′ can switch locations. In solution alternating purine–pyrimidine residues assume the left-handed conformation only in the presence of high concentrations of positively charged ions (e.g., Na^+) that shield the negatively charged phosphate groups. At lower salt concentrations, they form typical right-handed conformations. The physiological significance of Z DNA is uncertain and left-handed helices probably account at most for only a small of proportion of a cell's DNA. Further details of the A, B, and Z forms of DNA are presented in Table 6-2.

DNA Strands Can Separate (Denature) and Reassociate

Because the two strands of the double helix are held together by relatively weak (non-covalent) forces, you might expect that the two strands could come apart easily. Indeed, the original structure for the double

TABLE 6-2 A Comparison of the Structural Properties of A, B, and Z DNAs as Derived from Single-Crystal X-Ray Analysis

	Helix Type		
	A	**B**	**Z**
Overall proportions	Short and broad	Longer and thinner	Elongated and slim
Rise per base pair	2.3 Å	3.32 Å	3.8 Å
Helix-packing diameter	25.5 Å	23.7 Å	18.4 Å
Helix rotation sense	Right-handed	Right-handed	Left-handed
Base pairs per helix repeat	1	1	2
Base pairs per turn of helix	~11	~10	12
Rotation per base pair	33.6°	35.9°	−60° per 2 bp
Pitch per turn of helix	24.6 Å	33.2 Å	45.6 Å
Tilt of base normals to helix axis	+19°	−1.2°	−9°
Base-pair mean propeller twist	+18°	+16°	~0°
Helix axis location	Major groove	Through base pairs	Minor groove
Major-groove proportions	Extremely narrow but very deep	Wide and of intermediate depth	Flattened out on helix surface
Minor-groove proportions	Very broad but shallow	Narrow and of intermediate depth	Extremely narrow but very deep
Glycosyl-bond conformation	anti	anti	anti at C, syn at G

Source: Dickerson, R. E. et al. 1982. *Cold Spring Harbor Symp. Quant. Biol.* **47**:14. Reproduced by permission.

helix suggested that DNA replication would occur in just this manner. The complementary strands of double helix can also be made to come apart when a solution of DNA is heated above physiological temperatures (to near 100 °C) or under conditions of high pH, a process known as **denaturation.** However, this complete separation of DNA strands by denaturation is reversible. When heated solutions of denatured DNA are slowly cooled, single strands often meet their complementary strands and reform regular double helices (Figure 6-14). The capacity to renature denatured DNA molecules permits artificial hybrid DNA molecules to be formed by slowly cooling mixtures of denatured DNA from two different sources. Likewise, hybrids can be formed between complementary strands of DNA and RNA. As we shall see in Chapter 20, the ability to form hybrids between two single-stranded nucleic acids (**hybridization**) is the basis for several indispensable techniques in molecular biology, such as Southern blot hybridization and DNA microarrays.

Important insights into the properties of the double helix were obtained from classic experiments carried out in the 1950s in which the denaturation of DNA was studied under a variety of conditions. In these experiments DNA denaturation was monitored by measuring the absorbance of ultraviolet light passed through a solution of DNA. DNA maximally absorbs ultraviolet light at a wavelength of about 260 nm. It is the bases that are principally responsible for this absorption. When the temperature of a solution of DNA is raised to near the boiling point of water, the optical density (absorbance) at 260 nm markedly increases. The explanation for this increase is that duplex DNA is **hypochromic;** it absorbs less ultraviolet light by about 40% than do individual DNA chains. The hypochromicity is due to base stacking, which diminishes the capacity of the bases in duplex DNA to absorb ultraviolet light.

If we plot the optical density of DNA as a function of temperature, we observe that the increase in absorption occurs abruptly over a relatively narrow temperature range. The midpoint of this transition is the **melting point** or T_m (Figure 6-15). Like ice, DNA melts: it undergoes a transition

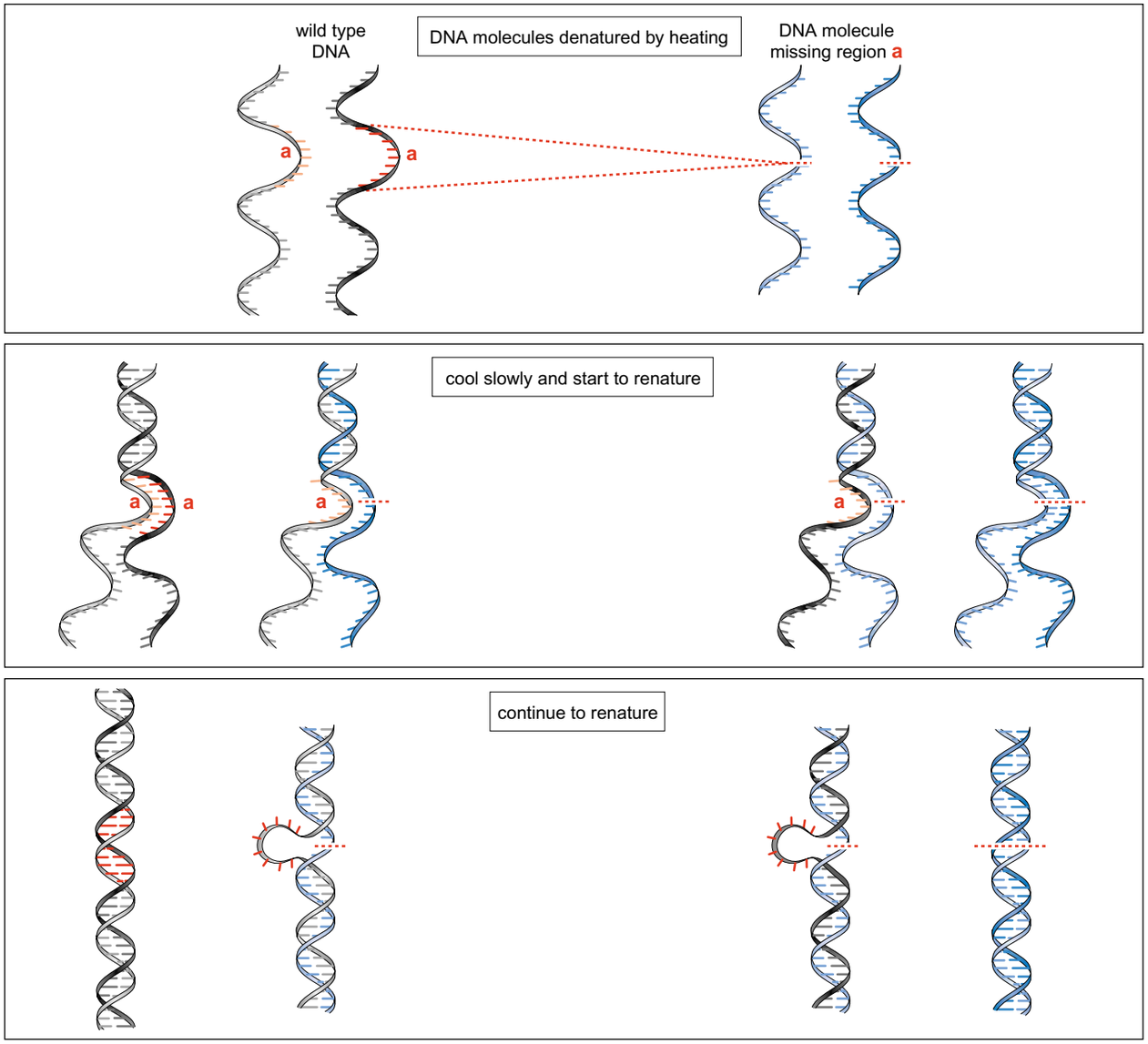


FIGURE 6-14 Reannealing and Hybridization. A mixture of two otherwise identical double-stranded DNA molecules, one normal wild type DNA and the other a mutant missing a short stretch of nucleotides (marked as region a in red), are denatured by heating. The denatured DNA molecules are allowed to renature by incubation just below the melting temperature. This treatment results in two types of renatured molecules. One type is composed of completely renatured molecules in which two complementary wild type strands reform a helix and two complementary mutant strands reform a helix. The other type are hybrid molecules, composed of a wild type and a mutant strand, exhibiting a short unpaired loop of DNA (region a).

from a highly ordered double-helical structure to a much less ordered structure of individual strands. The sharpness of the increase in absorbance at the melting temperature tells us that the denaturation and renaturation of complementary DNA strands is a highly cooperative, zippering-like process. Renaturation, for example, probably occurs by means of a slow nucleation process in which a relatively small stretch of bases on one strand find and pair with their complement on the complementary strand (middle panel of Figure 6-14). The remainder of the two strands then rapidly zipper-up from the nucleation site to reform an extended double helix (lower panel of Figure 6-14).

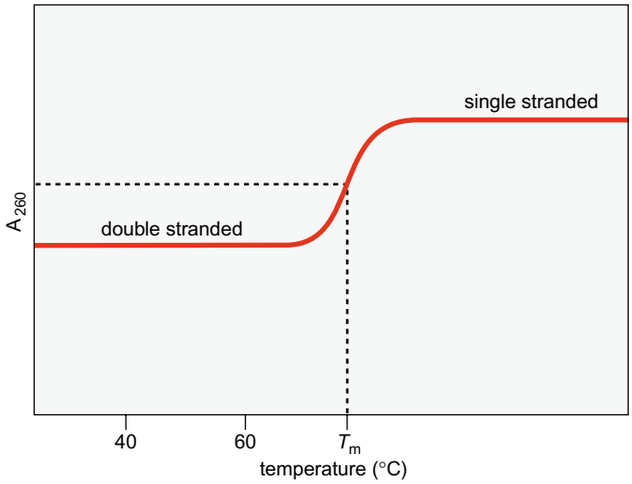


FIGURE 6-15 DNA Denaturation Curve.

The melting temperature of DNA is a characteristic of each DNA that is largely determined by the G:C content of the DNA and the ionic strength of the solution. The higher the percent of G:C base pairs in the DNA (and hence the lower the content of A:T base pairs), the higher the melting point (Figure 6-16). Likewise, the higher the salt concentration of the solution the greater the temperature at which the DNA denatures. How do we explain this behavior? G:C base pairs contribute more to the stability of DNA than do A:T base pairs because of the greater number of hydrogen bonds for the former (three in a G:C base pair versus two for A:T) but also importantly because the stacking interactions of G:C base pairs with adjacent base pairs are more favorable than the corresponding interactions of A:T base pairs with their neighboring base pairs. The effect of ionic strength reflects another fundamental feature of the double helix. The backbones of the two DNA strands contain phosphoryl groups, which carry a negative charge. These negative charges are close enough across the two strands that if not shielded they tend to cause the strands to repel each other, facilitating their separation. At high ionic strength, the negative charges are shielded by

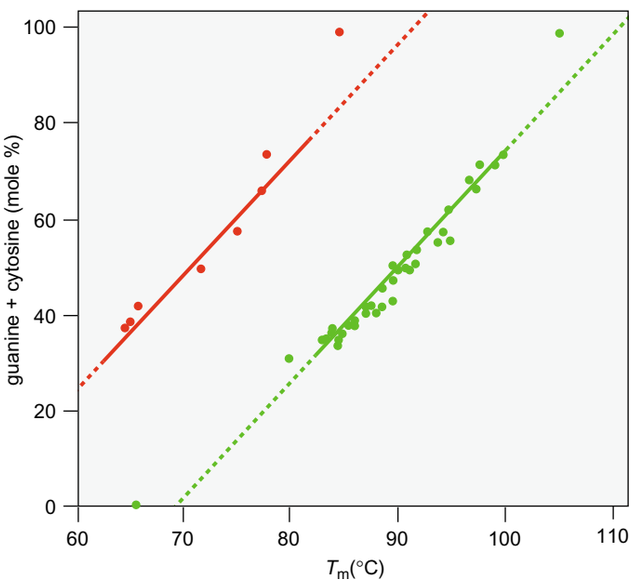


FIGURE 6-16 Dependence of DNA Denaturation on G + C Content and on Salt Concentration. The greater the G + C content, the higher the temperature must be to denature the DNA strand. DNA from different sources was dissolved in solutions of low (red line) and high (green line) concentrations of salt at pH 7.0. The points represent the temperature at which the DNA denatured, graphed against the G + C content. (Source: Data from Marmur, J. and Doty, P. 1962. *J. Mol. Biol.* **5**:120.)

cations, thereby stabilizing the helix. Conversely, at low ionic strength the unshielded negative charges render the helix less stable.

Some DNA Molecules Are Circles

It was initially believed that all DNA molecules are linear and have two free ends. Indeed, the chromosomes of eukaryotic cells each contain a single (extremely long) DNA molecule. But now we know that some DNAs are circles. For example, the chromosome of the small monkey DNA virus SV40 is a circular, double-helical DNA molecule of about 5,000 base pairs. Also, most (but not all) bacterial chromosomes are circular; *E. coli* has a circular chromosome of about 5 million base pairs. Additionally, many bacteria have small autonomously replicating genetic elements known as **plasmids,** which are generally circular DNA molecules.

Interestingly, some DNA molecules are sometimes linear and sometimes circular. The most well-known example is that of the bacteriophage λ, a DNA virus of *E. coli*. The phage λ genome is a linear double-stranded molecule in the virion particle. However, when the λ genome is injected into an *E. coli* cell during infection, the DNA circularizes. This occurs by base-pairing between single-stranded regions that protrude from the ends of the DNA and that have complementary sequences (“sticky ends”).

DNA TOPOLOGY

As DNA is a flexible structure, its exact molecular parameters are a function of both the surrounding ionic environment and the nature of the DNA-binding proteins with which it is complexed. Because their ends are free, linear DNA molecules can freely rotate to accommodate changes in the number of times the two chains of the double helix twist about each other. But if the two ends are covalently linked to form a circular DNA molecule and if there are no interruptions in the sugar phosphate backbones of the two strands, then the absolute number of times the chains can twist about each other cannot change. Such a covalently closed, circular DNA is said to be topologically constrained. Even the linear DNA molecules of eukaryotic chromosomes are subject to topological constraints due to their entrainment in chromatin and interaction with other cellular components (see Chapter 7). Despite these constraints, DNA participates in numerous dynamic processes in the cell. For example, the two strands of the double helix, which are twisted around each other, must rapidly separate in order for DNA to be duplicated and to be transcribed into RNA. Thus, understanding the topology of DNA and how the cell both accommodates and exploits topological constraints during DNA replication, transcription, and other chromosomal transactions is of fundamental importance in molecular biology.

Linking Number Is an Invariant Topological Property of Covalently Closed, Circular DNA

Let us consider the topological properties of **covalently closed, circular DNA,** which is referred to as **cccDNA.** Because there are no interruptions in either polynucleotide chain, the two strands of cccDNA cannot be separated from each other without the breaking of a covalent bond. If we wished to separate the two circular strands without

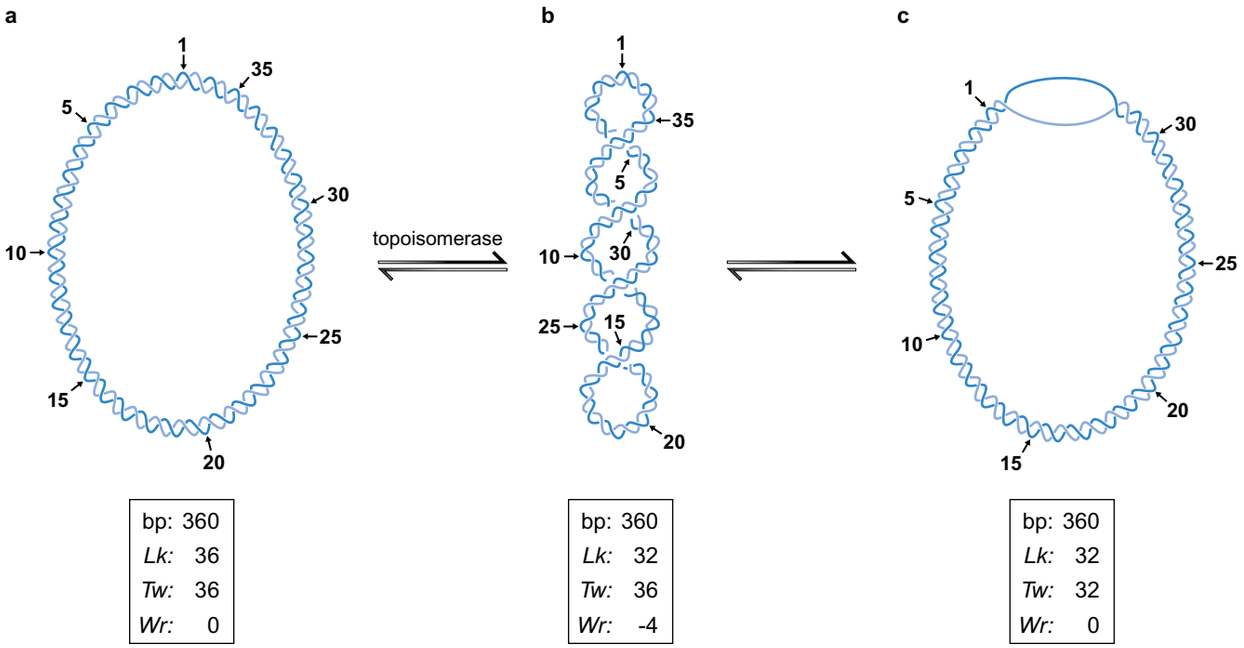


FIGURE 6-17 Topological States of Covalently Closed Circular (ccc) DNA. The figure shows conversion of the relaxed (a) to the negatively supercoiled (b) form of DNA. The strain in the supercoiled form may be taken up by supertwisting (b) or by local disruption of base pairing (c). [Adapted from a diagram provided by Dr. M. Gellert.] (Source: Modified from Kornberg, A. and Baker, T. A. 1992. *DNA Replication*. Figure 1-21, page 32)

permanently breaking any bonds in the sugar phosphate backbones, we would have to pass one strand through the other strand repeatedly (we will encounter an enzyme that can perform just this feat!). The number of times one strand would have to be passed through the other strand in order for the two strands to be entirely separated from each other is called the **linking number** (Figure 6-17). The linking number, which is always an integer, is an invariant topological property of cccDNA, no matter how much the DNA molecule is distorted.

Linking Number Is Composed of Twist and Writhe

The linking number is the sum of two geometric components called the **twist** and the **writhe.** Let us consider twist first. Twist is simply the number of helical turns of one strand about the other, that is, the number of times one strand completely wraps around the other strand. Consider a cccDNA that is lying flat on a plane. In this flat conformation, the linking number is fully composed of twist. Indeed, the twist can be easily determined by counting the number of times the two strands cross each other (see Figure 6-17a). The helical crossovers (twist) in a right-handed helix are defined as positive such that the linking number of DNA will have a positive value.

But cccDNA is generally not lying flat on a plane. Rather, it is usually torsionally stressed such that the long axis of the double helix crosses over itself, often repeatedly, in three-dimensional space. This is called *writhe*. To visualize the distortions caused by torsional stress, think of the coiling of a telephone cord that has been overtwisted (Figure 6-17b).

Writhe can take two forms. One form is the **interwound or plectonemic writhe,** in which the long axis is twisted around itself, as depicted in Figure 6-17b and Figure 6-18a. The other form of writhe is

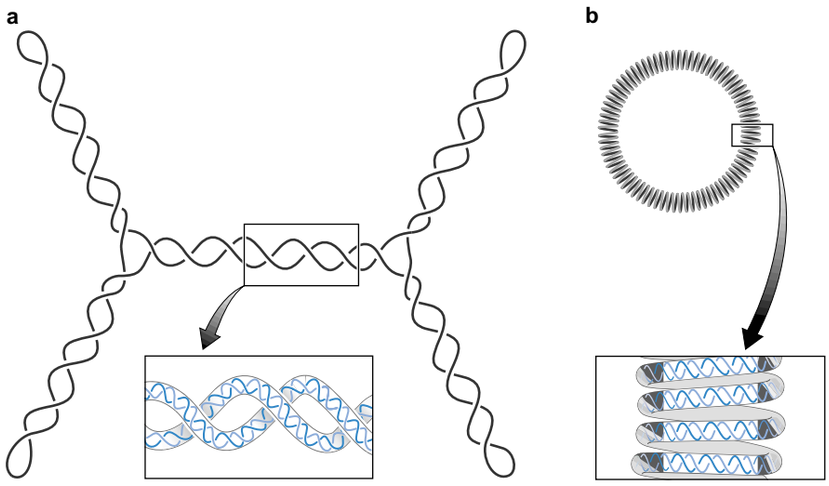


FIGURE 6-18 Two Forms of Writhe of Supercoiled DNA. The figure shows interwound (a) and toroidal (b) writhe of cccDNA of the same length. (a) The interwound or plectonemic writhe is formed by twisting of the double helical DNA molecule over itself as depicted in the example of a branched molecule. (b) Toroidal or spiral writhe is depicted in this example by cylindrical coils. [Courtesy of Dr. N. R. Cozzarelli] (Source: Adapted from Kornberg, A. and Baker, T. A. 1992. *DNA Replication*. Figure 1-22, page 33)

a **toroid** or **spiral** in which the long axis is wound in a cylindrical manner, as often occurs when DNA wraps around protein (Figure 6-18b). The **writhing number** (***Wr***) is the total number of interwound and/or spiral writhes in cccDNA. For example, the molecule shown in Figure 6-17b has a writhe of 4 from 4 interwound writhes.

Interwound writhe and spiral writhe are topologically equivalent to each other and are readily interconvertible geometric properties of cccDNA. Also, twist and writhe are interconvertible. A molecule of cccDNA can readily undergo distortions that convert some of its twist to writhe or some of its writhe to twist without the breakage of any covalent bonds. The only constraint is that the sum of the **twist number** (***Tw***) and the writhing number (*Wr*) must remain equal to the **linking number** (***Lk***). This constraint is described by the equation: $Lk = Tw + Wr$.

Lk^O Is the Linking Number of Fully Relaxed cccDNA under Physiological Conditions

Consider cccDNA that is free of **supercoiling** (that is, it is said to be **relaxed**) and whose twist corresponds to that of the B form of DNA in solution under physiological conditions (about 10.5 base pairs per turn of the helix). The linking number (*Lk*) of such cccDNA under physiological conditions is assigned the symbol $\boldsymbol{Lk^O}$**.** Lk^O for such a molecule is the number of base pairs divided by 10.5. For a cccDNA of 10,500 base pairs, $Lk = +1{,}000$. (The sign is positive because the twists of DNA are right-handed.) One way to see this is to imagine pulling one strand of the 10,500 base pair cccDNA out into a flat circle. If we did this, then the other strand would cross the flat circular strand 1,000 times.

How can we remove supercoils from cccDNA if it is not already relaxed? One procedure is to treat the DNA mildly with the enzyme DNase I, so as to break on average one phosphodiester bond (or a small number of bonds) in each DNA molecule. Once the DNA has been "nicked" in this manner, it is no longer topologically constrained and the strands can rotate freely, allowing writhe to dissipate (Figure 6-19). If the nick is then repaired, the resulting cccDNA molecules will be relaxed and will have on average an *Lk* that is equal to Lk^O. (Due to rotational fluctuation at the time the nick is repaired, some of

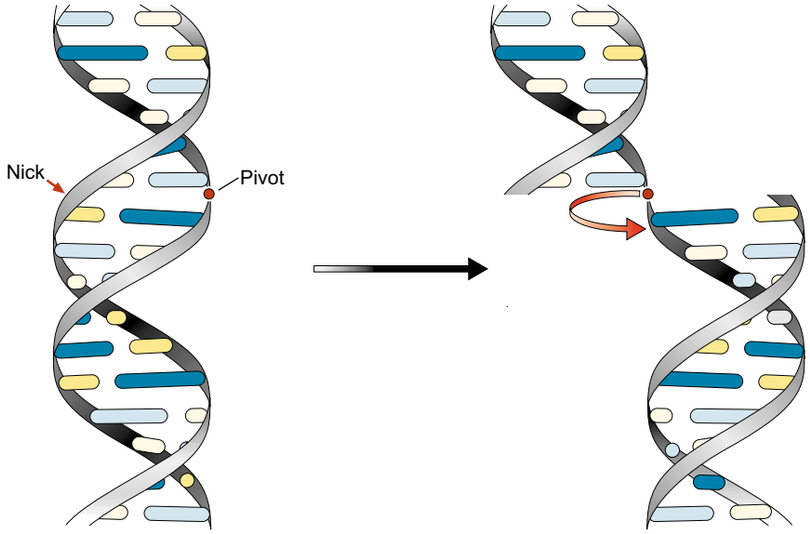


FIGURE 6-19 Relaxing DNA with DNase I.

the resulting cccDNAs will have an *Lk* that is somewhat greater than Lk^O and others will have an *Lk* that is somewhat lower. Thus, the relaxation procedure will generate a narrow spectrum of topoisomers whose average *Lk* is equal to Lk^O).

DNA in Cells Is Negatively Supercoiled

The extent of supercoiling is measured by the difference between *Lk* and Lk^O, which is called the **linking difference:**

$$\Delta Lk = Lk - Lk^O.$$

If the ΔLk of a cccDNA is significantly different from zero, then the DNA is torsionally strained and hence it is supercoiled. If $Lk < Lk^O$ and $\Delta Lk < 0$, then the DNA is said to be "negatively supercoiled." Conversely, if $Lk > Lk^O$ and $\Delta Lk > 0$, then the DNA is "positively supercoiled." For example, the molecule shown in Figure 6-17b is negatively supercoiled and has a linking difference of −4 because its *Lk* (32) is four less than that (36) for the relaxed form of the molecule shown in Figure 6-17a.

Because ΔLk and Lk^O are dependent upon the length of the DNA molecule, it is more convenient to refer to a normalized measure of supercoiling. This is the **superhelical density,** which is assigned the symbol σ and is defined as:

$$\sigma = \Delta Lk/Lk^O$$

DNA rings purified both from bacteria and eukaryotes are usually negatively supercoiled, having values of σ of about −0.06. The electron micrograph shown in Figure 6-20 compares the structures of bacteriophage DNA in its relaxed form with its supercoiled form.

What does this mean biologically? Negative supercoils can be thought of as a store of free energy that aids in processes that require strand separation, such as DNA replication and transcription. Because $Lk = Tw + Wr$, negative supercoils can be converted into untwisting of the double helix (compare Figure 6-17a with 6-17b). Regions of negatively supercoiled DNA therefore have a tendency to partially unwind. Thus, strand

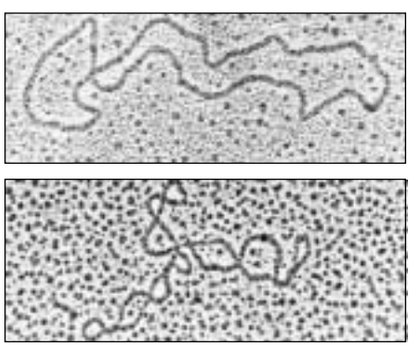


FIGURE 6-20 EM of Supercoiled DNA. The upper electron micrograph is a relaxed (nonsupercoiled) DNA molecule of bacteriophage PM2. The lower electron micrograph shows the phage in its supertwisted form. (Source: Electron micrographs courtesy of Wang, J. C. 1982. *Sci. Am.* **247**:97.)

separation can be accomplished more easily in negatively supercoiled DNA than in relaxed DNA.

The only organisms that have been found to have positively supercoiled DNA are certain thermophiles, microorganisms that live under conditions of extreme high temperatures, such as in hot springs. In this case, the positive supercoils can be thought of as a store of free energy that helps keep the DNA from denaturing at the elevated temperatures. In so far as positive supercoils can be converted into more twist (positively supercoiled DNA can be thought of as being overwound), strand separation requires more energy in thermophiles than in organisms whose DNA is negatively supercoiled.

Nucleosomes Introduce Negative Supercoiling in Eukaryotes

As we shall see in the next chapter, DNA in the nucleus of eukaryotic cells is packaged in small particles known as **nucleosomes** in which the double helix is wrapped almost two times around the outside circumference of a protein core. You will be able to recognize this wrapping as the toroid or spiral form of writhe. Importantly, it occurs in a left-handed manner. (Convince yourself of this by applying the handedness rule in your mind's eye to DNA wrapped around the nucleosome in Chapter 7, Figure 7-8). It turns out that writhe in the form of left-handed spirals is equivalent to negative supercoils. Thus, the packaging of DNA into nucleosomes introduces negative superhelical density.

Topoisomerases Can Relax Supercoiled DNA

As we have seen, the linking number is an invariant property of DNA that is topologically constrained. It can only be changed by introducing interruptions into the sugar-phosphate backbone. A remarkable class of enzymes known as **topoisomerases** are able to do just that by introducing transient nicks or breaks into the DNA. Topoisomerases are of two broad types. Type II topoisomerases change the linking number in steps of two. They make transient double-stranded breaks in the DNA, through which they pass a region of uncut duplex DNA before resealing the break (Figure 6-21). Type II topoisomerases require energy from ATP hydrolysis for their action. Type I topoisomerases, in contrast, change

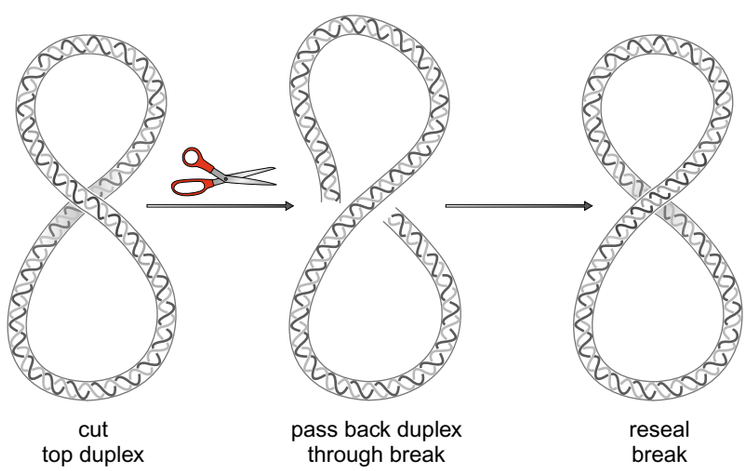


FIGURE 6-21 Schematic for Changing the Linking Number in DNA with Topoisomerase II. Topoisomerase II binds to DNA, creates a double-stranded break, passes uncut DNA through the gap, then reseals the break.

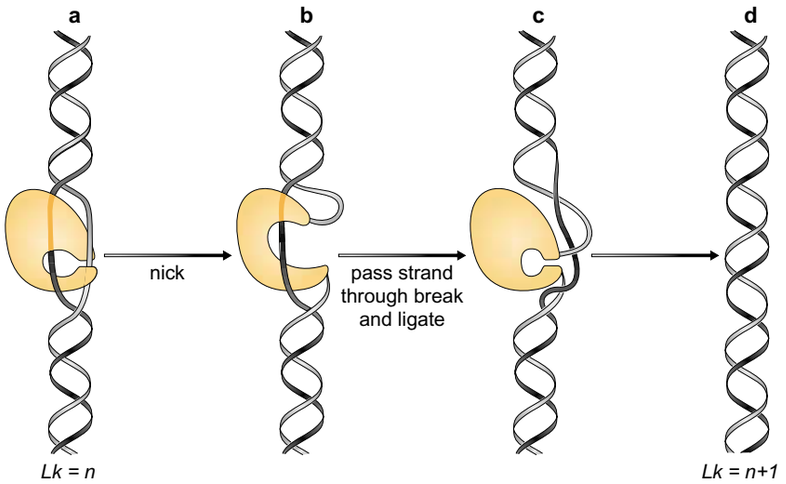


FIGURE 6-22 Schematic Mechanism of Action for Topoisomerase I. (a) The enzyme binds to DNA. (b) It then nicks one strand and prevents the free rotation of the helix by remaining bound to each broken end. (c) The enzyme passes the other strand through the break and ligates the cut ends, thereby increasing the linking number of the DNA by 1. (d) The enzyme falls away and the strands renature, leaving a DNA with the linking number increasing by 1. (Source: Redrawn from Dean, F. et al. 1983. *Cold Spring Harbor Symp. Quant. Biol.* **47**:773.)

the linking number of DNA in steps of one. They make transient single-stranded breaks in the DNA, allowing one strand to pass through the break in the other before resealing the nick (Figure 6-22). Type I topoisomerases relax DNA by removing supercoils (dissipating writhe). They can be compared to the protocol of introducing nicks into cccDNA with DNase and then repairing the nicks, which as we saw can be used to relax cccDNA, except that type I topoisomerases relax DNA in a controlled and concerted manner (Figure 6-22). In contrast to type II topoisomerases, type I topoisomerases do not require ATP. As we shall see in Chapter 10, both type I and type II topoisomerases work through an intermediate in which the enzyme is covalently attached to one end of the broken DNA.

Prokaryotes Have a Special Topoisomerase That Introduces Supercoils

Both prokaryotes and eukaryotes have type I and type II topoisomerases, which are capable of removing supercoils from DNA. In addition, however, prokaryotes have a special type II topoisomerase known as **DNA gyrase** that is able to introduce negative supercoils, rather than remove them. DNA gyrase is responsible for the negative supercoiling of chromosomes in prokaryotes, which facilitates unwinding of the DNA duplex during transcription and DNA replication.

DNA Topoisomers Can Be Separated by Electrophoresis

Covalently closed, circular DNA molecules of the same length but of different linking numbers are called **DNA topoisomers.** Even though topoisomers have the same molecular weight, they can be separated from each other by electrophoresis through a gel of agarose (see Chapter 20 for an explanation of **gel electrophoresis**). The basis for this separation is that the greater the writhe the more compact the shape of a cccDNA. Once again, think of how supercoiling a telephone cord causes it to become more compact. The more compact the DNA, the more easily (up to a point) it is able to migrate through the gel matrix (Figure 6-23). Thus, a fully relaxed cccDNA migrates more slowly than a highly supercoiled topoisomer of the same circular DNA. Figure 6-24 shows a ladder

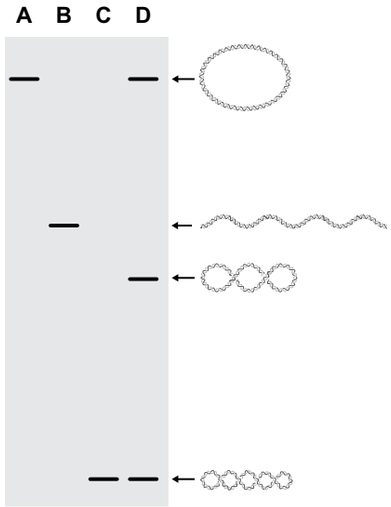


FIGURE 6-23 Schematic of Electrophoretic Separation of DNA Topoisomers. Lane A represents relaxed or nicked circular DNA; lane B, linear DNA; lane C, highly supercoiled ccDNA; and lane D, a ladder of topoisomers.

of DNA topoisomers resolved by gel electrophoresis. Molecules in adjacent rungs of the ladder differ from each other by a linking number difference of just one. Obviously, electrophoretic mobility is highly sensitive to the topological state of DNA (see Box 6-2, below).

Box 6-2 Proving That DNA Has a Helical Periodicity of about 10.5 Base Pairs per Turn from the Topological Properties of DNA Rings

The observation that DNA topoisomers can be separated from each other electrophoretically is the basis for a simple experiment that proves that DNA has a helical periodicity of about 10.5 base pairs per turn in solution. Consider three cccDNAs of sizes 3990, 3995, and 4011 base pairs that were relaxed to completion by treatment with topoisomerase I. When subjected to electrophoresis through agarose, the 3990- and 4011-base-pair DNAs exhibit essentially identical mobilities. Due to thermal fluctuation, topoisomerase treatment actually generates a narrow spectrum of topoisomers, but for simplicity let us consider the mobility of only the most abundant topoisomer (that corresponding to the cccDNA in its most relaxed state). The mobilities of the most abundant topoisomers for the 3990- and 4011-base-pair DNAs are indistinguishable because the 21-base-pair difference between them is negligible compared to the sizes of the rings. The most abundant topoisomer for the 3995-base-pair ring, however, is found to migrate slightly more rapidly than the other two rings even though it is only 5 base pairs larger than the 3990-base-pair ring. How are we to explain this anomaly? The 3990- and 4011-base-pair rings in their most relaxed states are expected to have linking numbers equal to Lk^O, that is, 380 in the case of the 3990-base-pair ring (dividing the size by 10.5 base pairs) and 382 in the case of the 4011-base-pair ring. Because Lk is equal to Lk^O, the linking difference ($\Delta Lk = Lk - Lk^O$) in both cases is zero and there is no writhe. But because the linking number must be an integer, the most relaxed state for the 3995-base-pair ring would be either of two topoisomers having linking numbers of 380 or 381. However, Lk^O for the 3995-base-pair ring is 380.5. Thus, even in its most relaxed state, a covalently closed circle of 3995 base pairs would necessarily have about half a unit of writhe (its linking difference would be 0.5), and hence it would migrate more rapidly than the 3990- and 4011-base-pair circles. In other words, to explain how rings that differ in length by 21 base pairs (two turns of the helix) have the same mobility whereas a ring that differs in length by only 5 base pairs (about half a helical turn) exhibits a different mobility, we must conclude that DNA in solution has a helical periodicity of about 10.5 base pairs per turn.

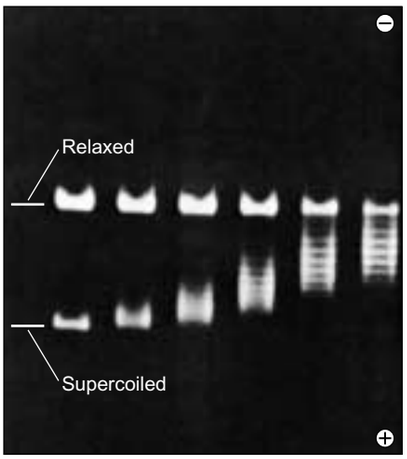


FIGURE 6-24 Separation of Relaxed and Supercoiled DNA by Gel Electrophoresis. Relaxed and supercoiled DNA topoisomers are resolved by gel electrophoresis. The speed with which the DNA molecules migrate increases as the number of superhelical turns increases. (Source: Courtesy of J. C. Wang.)

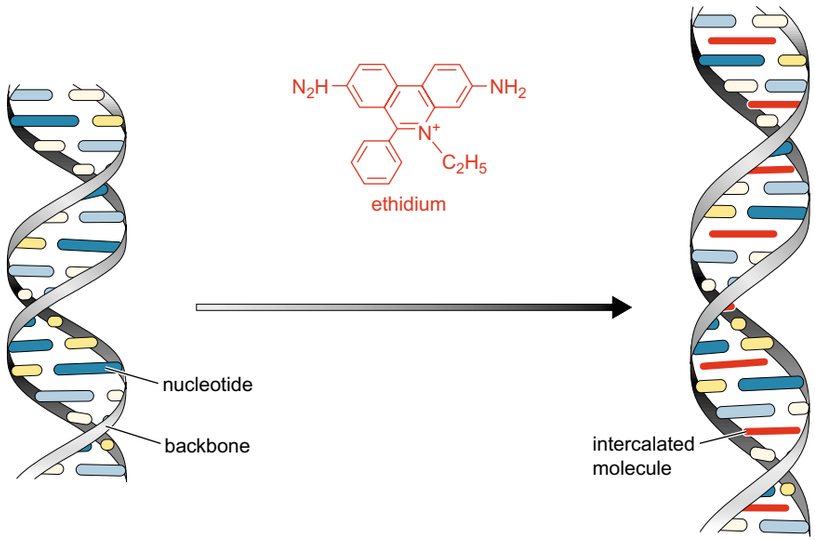


FIGURE 6-25 Intercalation of Ethidium Bromide into DNA. Ethidium bromide increases the spacing of successive base pairs, distorts the regular sugar-phosphate backbone, and increases the pitch of the helix.

Ethidium Ions Cause DNA to Unwind

Ethidium is a large, flat, multi-ringed cation. Its planar shape enables ethidium to slip (intercalate) between the stacked base pairs of DNA (Figure 6-25). Because it fluoresces when exposed to ultraviolet light, and because its fluorescence increases dramatically after intercalation, ethidium is used as a stain to visualize DNA.

When an ethidium ion intercalates between two base pairs, it causes the DNA to unwind by 26°, reducing the normal rotation per base pair from ~36° to ~10°. In other words, ethidium decreases the twist of DNA. Imagine the extreme case of a DNA molecule that has an ethidium ion between every base pair. Instead of 10 base pairs per turn it would have 36! When ethidium binds to linear DNA or to a nicked circle, it simply causes the helical pitch to increase. But consider what happens when ethidium binds to covalently closed, circular DNA. The linking number of the cccDNA does not change (no covalent bonds are broken and resealed), but the twist decreases by 26° for each molecule of ethidium that has bound to the DNA. Because $Lk = Tw + Wr$, this decrease in *Tw* must be compensated for by a corresponding increase in *Wr*. If the circular DNA is initially negatively supercoiled (as is normally the case for circular DNAs isolated from cells), then the addition of ethidium will increase *Wr*. In other words, the addition of ethidium will relax the DNA. If enough ethidium is added, the negative supercoiling will be brought to zero, and if even more ethidium is added, *Wr* will increase above zero and the DNA will become positively supercoiled.

Because the binding of ethidium increases *Wr*, its presence greatly affects the migration of cccDNA during gel electrophoresis. In the presence of non-saturating amounts of ethidium, negatively supercoiled circular DNAs are more relaxed and migrate more slowly, whereas relaxed cccDNAs become positively supercoiled and migrate more rapidly.

RNA STRUCTURE

RNA Contains Ribose and Uracil and Is Usually Single-Stranded

We now turn our attention to RNA, which differs from DNA in three respects (Figure 6-26). First, the backbone of RNA contains ribose rather than 2′-deoxyribose. That is, ribose has a hydroxyl group at the 2′ position. Second, RNA contains **uracil** in place of thymine. Uracil has the same single-ringed structure as thymine, except that it lacks the 5′ methyl group. Thymine is in effect 5′methyl-uracil. Third, RNA is usually found as a single polynucleotide chain. Except for the case of certain viruses, RNA is not the genetic material and does not need to be capable of serving as a template for its own replication. Rather, RNA functions as the intermediate, the mRNA, between the gene and the protein-synthesizing machinery. Another function of RNA is as an adaptor, the tRNA, between the codons in the mRNA and amino acids. RNA can also play a structural role as in the case of the RNA components of the ribosome. Yet another role for RNA is as a regulatory molecule, which through sequence complementarity binds to, and interferes with the translation of, certain mRNAs. Finally, some RNAs (including one of the structural RNAs of the ribosome) are enzymes that catalyze essential reactions in the cell. In all of these cases, the RNA is copied as a single strand off only one of the two strands of the DNA template, and its complementary strand does not exist. RNA is capable of forming long double helices, but these are unusual in nature.

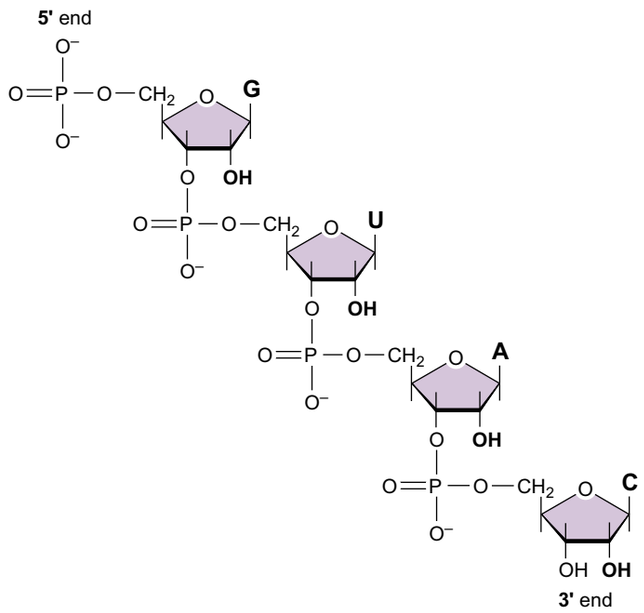


FIGURE 6-26 Structural Features of RNA. The figure shows the structure of the backbone of RNA, composed of alternating phosphate and ribose moieties.

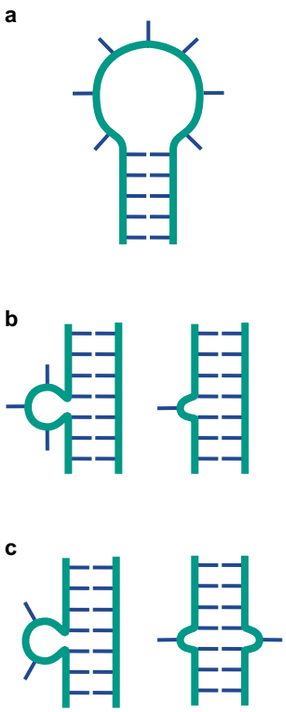


FIGURE 6-27 Double Helical Characteristics of RNA. In an RNA molecule having regions of complementary sequences, the intervening (non-complementary) stretches of RNA may become "looped out" to form one of the structures illustrated in the figure.
(a) Hairpin (b) Bulge (c) Loop

RNA Chains Fold Back on Themselves to Form Local Regions of Double Helix Similar to A-Form DNA

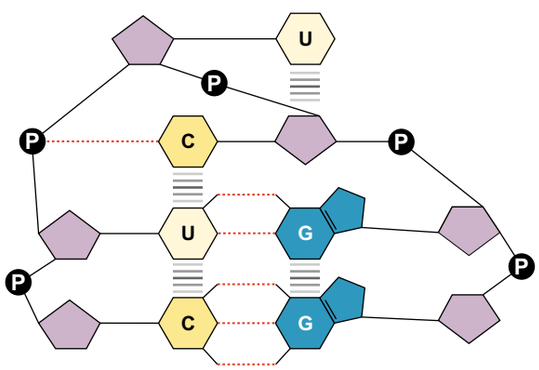
Despite being single-stranded, RNA molecules often exhibit a great deal of double-helical character (Figure 6-27). This is because RNA chains frequently fold back on themselves to form base-paired segments between short stretches of complementary sequences. If the two stretches of complementary sequence are near each other, the RNA may adopt one of various **stem-loop structures** in which the intervening RNA is looped out from the end of the double-helical segment as in a hairpin, a bulge, or a simple loop.

The stability of such stem-loop structures is in some instances enhanced by the special properties of the loop. For example, a stem-loop with the "tetraloop" sequence UUCG is unexpectedly stable due to special base-stacking interactions in the loop (Figure 6-28). Base pairing can also take place between sequences that are not contiguous to form complex structures aptly named **pseudoknots** (Figure 6-29). The regions of base pairing in RNA can be a regular double helix or they can contain discontinuities, such as noncomplementary nucleotides that bulge out from the helix.

A feature of RNA that adds to its propensity to form double-helical structures is an additional, non-Watson-Crick base pair. This is the G:U base pair, which has hydrogen bonds between N3 of uracil and the carbonyl on C6 of guanine and between the carbonyl on C2 of uracil and N1 of guanine (Figure 6-30). Because G:U base pairs can occur as well as the four conventional, Watson-Crick base pairs, RNA chains have an enhanced capacity for self-complementarity. Thus, RNA frequently exhibits local regions of base pairing but not the long-range, regular helicity of DNA.

The presence of 2′-hydroxyls in the RNA backbone prevents RNA from adopting a B-form helix. Rather, double-helical RNA resembles the A-form structure of DNA. As such, the minor groove is wide and shallow, and hence accessible, but recall that the minor groove offers little sequence-specific information. Meanwhile, the major groove is so narrow and deep that it is not very accessible to amino acid side chains from interacting proteins. Thus, the RNA double helix is quite distinct from the DNA double helix in its detailed atomic structure and less well suited for sequence-specific interactions with proteins (although some proteins do bind to RNA in a sequence-specific manner).

FIGURE 6-28 Tetraloop. Base stacking interactions promote and stabilize the tetraloop structure. The black circles between the riboses represent the phosphate moieties of the RNA backbone.



C(UUCG)G Tetraloop

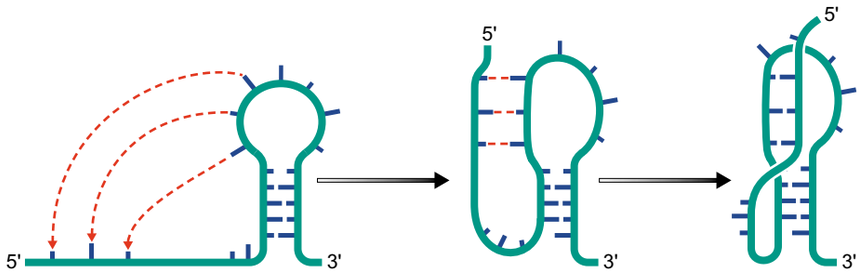


FIGURE 6-29 Pseudoknot. The pseudoknot structure is formed by base pairing between noncontiguous complementary sequences.

RNA Can Fold Up into Complex Tertiary Structures

Freed of the constraint of forming long-range regular helices, RNA can adopt a wealth of tertiary structures. This is because RNA has enormous rotational freedom in the backbone of its non-base-paired regions. Thus, RNA can fold up into complex tertiary structures frequently involving unconventional base pairing, such as the base triples and base-backbone interactions seen in tRNAs (see, for example, the illustration of the U:A:U base triple in Figure 6-31). Proteins can assist the formation of tertiary structures by large RNA molecules, such as those found in the ribosome. Proteins shield the negative charges of backbone phosphates, whose electrostatic repulsive forces would otherwise destabilize the structure.

Researchers have taken advantage of the potential structural complexity of RNA to generate novel RNA species (not found in nature) that have specific desirable properties. By synthesizing RNA molecules with randomized sequences, it is possible to generate mixtures of oligonucleotides representing enormous sequence diversity. For example, a mixture of oligoribonucleotides of length 20 and having four possible nucleotides at each position would have a potential complexity of 4^{20} sequences or 10^{12} sequences! From mixtures of diverse oligoribonucleotides, RNA molecules can be selected biochemically that have particular properties, such as an affinity for a specific small molecule.

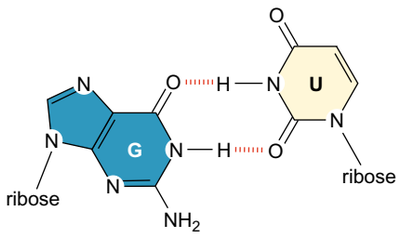
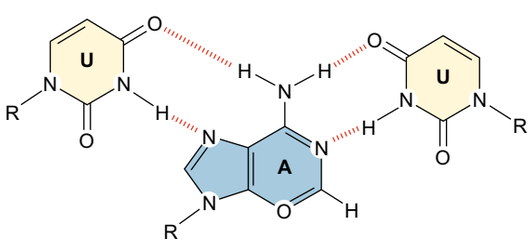


FIGURE 6-30 G:U Base Pair. The structure shows hydrogen bonds that allow base pairing to occur between guanine and uracil.

Some RNAs Are Enzymes

It was widely believed for many years that only proteins could be enzymes. An enzyme must be able to bind a substrate, carry out a chemical reaction, release the product and repeat this sequence of events many times. Proteins are well suited to this task because they are composed of many different kinds of amino acids (20) and they can fold into complex tertiary structures with binding pockets for the substrate and



U:A:U base triple

FIGURE 6-31 U:A:U Base Triple. The structure shows one example of hydrogen bonding that allows unusual triple base pairing.

small molecule cofactors and an active site for catalysis. Now we know that RNAs, which as we have seen can similarly adopt complex tertiary structures, can also be biological catalysts. Such RNA enzymes are known as **ribozymes,** and they exhibit many of the features of a classical enzyme, such as an active site, a binding site for a substrate and a binding site for a cofactor, such as a metal ion.

One of the first ribozymes to be discovered was **RNase P,** a ribonuclease that is involved in generating tRNA molecules from larger, precursor RNAs. RNase P is composed of both RNA and protein; however, the RNA moiety alone is the catalyst. The protein moiety of RNase P facilitates the reaction by shielding the negative charges on the RNA so that it can bind effectively to its negatively charged substrate. The RNA moiety is able to catalyze cleavage of the tRNA precursor in the absence of the protein if a small, positively charged counter ion, such as the peptide spermidine, is used to shield the repulsive, negative charges. Other ribozymes carry out trans-esterification reactions involved in the removal of intervening sequences known as **introns** from precursors to certain mRNAs, tRNAs, and ribosomal RNAs in a process known as **RNA splicing** (see Chapter 13).

The Hammerhead Ribozyme Cleaves RNA by the Formation of a 2′, 3′ Cyclic Phosphate

Before concluding our discussion of RNA, let us look in more detail at the structure and function of one particular ribozyme, the **hammerhead.** The hammerhead is a sequence-specific ribonuclease that is found in

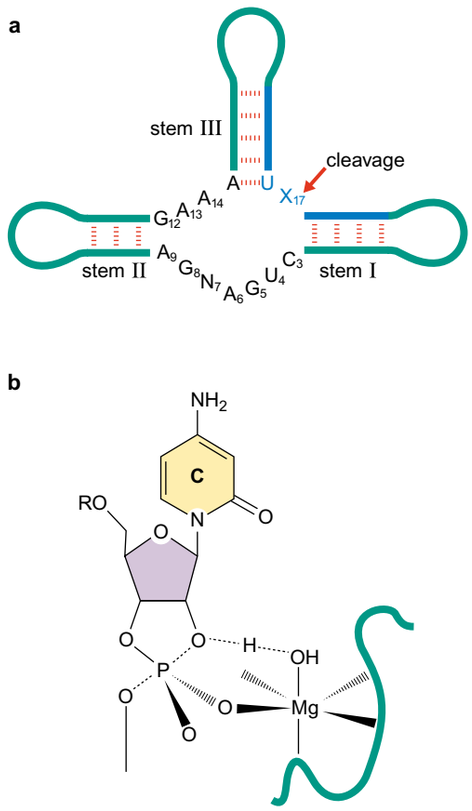


FIGURE 6-32 Hammerhead Ribozyme (Secondary Structure). (a) The figure shows the predicted secondary structures of the hammerhead ribozyme. Watson-Crick base-pair interactions are shown in red; the scissile bonds are shown by a red arrow; approximate minimal substrate strands are labeled in blue; (U) Uracil; (A) adenine; (C) cytosine; (G) guanine. (Source: Redrawn from McKay, D. B. and Wedekind, J. E. 1999. In *The RNA World,* 2nd edition (ed. Gesteland, R. F. et al.) Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press. Figure 1, part A, p. 267.) (b) The hammerhead ribozyme cleavage reaction involves an intermediary state during which Mg(OH) in complex with the ribozyme (shown in green) acts as a general base catalyst to remove a proton from the 2′ hydroxyl of the active site of cytosine (shown at position 17 in part (a)), and to initiate the cleavage reaction at the scissile phosphodiester bond at the active site. (Source: Redrawn from Scott, W. G. et al. 1995. *Cell* **81**:99; Figure 1, part B, p. 992.)

certain infectious RNA agents of plants known as *viroids*, which depend on self-cleavage to propagate. When the viroid replicates, it produces multiple copies of itself in one continuous RNA chain. Single viroids arise by cleavage, and this cleavage reaction is carried out by the RNA sequence around the junction. One such self-cleaving sequence is called the *hammerhead* because of the shape of its secondary structure, which consists of three base-paired stems (I, II, and III) surrounding a core of non-complementary nucleotides required for catalysis (Figure 6-32). The tertiary structure of the ribozyme, however, looks more like a wishbone (Figure 6-33).

To understand how the hammerhead works, let us first look at how RNA undergoes hydrolysis under alkaline conditions. At high pH, the 2′ hydroxyl of the ribose in the RNA backbone can become deprotonated, and the resulting negatively charged oxygen can attack the scissile phosphate at the 3′ position of the same ribose (Figure 6-32b). This reaction breaks the RNA chain, producing a 2′, 3′ cyclic phosphate and a free 5′ hydroxyl. Each ribose in an RNA chain can undergo this reaction, completely cleaving the parent molecule into nucleotides. (Why is DNA not similarly susceptible to alkaline hydrolysis?) Many protein ribonucleases also cleave their RNA substrates via the formation of a 2′, 3′ cyclic phosphate. Working at normal cellular pH, these protein enzymes use a metal ion, bound at their active site, to activate the 2′ hydroxyl of the RNA. The hammerhead is a sequence-specific ribonuclease, but it too cleaves RNA via the formation of a 2′, 3′ cyclic phosphate. Hammerhead-mediated cleavage involves a ribozyme-bound Mg^{++} ion that deprotonates the 2′ hydroxyl at neutral pH, resulting in **nucleophilic attack** on the scissile phosphate.

Because the normal reaction of the hammerhead is self-cleavage, it is not really a catalyst; each molecule normally promotes a reaction one time only, thus having a turnover number of one. But the hammerhead can be engineered to function as a true ribozyme by dividing the molecule into two portions—one, the ribozyme, that contains the catalytic core and the other, the substrate, that contains the cleavage site. The substrate binds to the ribozyme at stems I and III (Figure 6-33a). After cleavage, the substrate is released and replaced by a fresh uncut substrate, thereby allowing repeated rounds of cleavage.

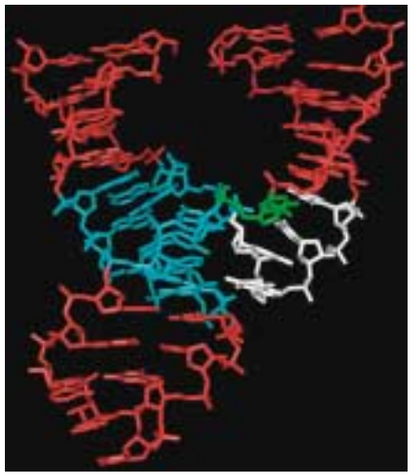


FIGURE 6-33 Hammerhead Ribozyme (Tertiary Structure). This view of the refined hammerhead ribozyme structure shows the conserved bases of stem III as well as the 3 bp augmenting helix that joins stem II (top left) to stem–loop III (bottom) highlighted in cyan, the CUGA uridine turn highlighted in white, and the active site cytosine (cut site at position 17) in green. The other helical residues are all shown in red to deemphasize the arbitrary distinction between enzyme and substrate strands. (Source: Scott, W. G., Finch, J. T., and Klug, A. 1995. *Cell* **81**:993.)

Did Life Evolve from an RNA World?

The discovery of ribozymes has profoundly altered our view of how life might have evolved. We can now imagine that there was a primitive form of life based entirely on RNA. In this world, RNA would have functioned as the genetic material and as the enzymatic machines. This RNA world would have preceded life as we know it today, in which information transfer is based on DNA, RNA, and protein. A hint that the protein world might have arisen from an RNA world is the discovery that the component in the ribosome that is responsible for the formation of the peptide bond, the peptidyl transferase, is an RNA molecule (see Chapter 14). Unlike RNase P, the hammerhead, and other previously known ribozymes which act on phosphorous centers, the peptidyl transferase acts on a carbon center to create the peptide bond. It thus links RNA chemistry to the most fundamental reaction in the protein world, peptide bond formation. Perhaps then the ribosome ribozyme is a relic of an earlier form of life in which all enzymes were RNAs.

SUMMARY

DNA is usually in the form of a right-handed double helix. The helix consists of two polydeoxynucleotide chains. Each chain is an alternating polymer of deoxyribose sugars and phosphates that are joined together via phosphodiester linkages. One of four bases protrudes from each sugar: adenine and guanine, which are purines, and thymine and cytosine, which are pyrimidines. While the sugar phosphate backbone is regular, the order of bases is irregular and this is responsible for the information content of DNA. Each chain has a 5′ to 3′ polarity, and the two chains of the double helix are oriented in an antiparallel manner—that is, they run in opposite directions.

Pairing between the bases holds the chains together. Pairing is mediated by hydrogen bonds and is specific: Adenine on one chain is always paired with thymine on the other chain, whereas guanine is always paired with cytosine. This strict base-pairing reflects the fixed locations of hydrogen atoms in the purine and pyrimidine bases in the forms of those bases found in DNA. Adenine and cytosine almost always exist in the amino as opposed to the imino tautomeric forms, whereas guanine and thymine almost always exist in the keto as opposed to enol forms. The complementarity between the bases on the two strands gives DNA its self-coding character.

The two strands of the double helix fall apart (denature) upon exposure to high temperature, extremes of pH, or any agent that causes the breakage of hydrogen bonds. Upon slow return to normal cellular conditions, the denatured single strands can specifically reassociate to biologically active double helices (renature or anneal).

DNA in solution has a helical periodicity of about 10.5 base pairs per turn of the helix. The stacking of base pairs upon each other creates a helix with two grooves. Because the sugars protrude from the bases at an angle of about 120°, the grooves are unequal in size. The edges of each base pair are exposed in the grooves, creating a pattern of hydrogen bond donors and acceptors and of van der Waals surfaces that identifies the base pair. The wider—or *major*—groove is richer in chemical information than the narrow (*minor*) groove and is more important for recognition by nucleotide sequence-specific binding proteins.

Almost all cellular DNAs are extremely long molecules, with only one DNA molecule within a given chromosome. Eukaryotic cells accommodate this extreme length in part by wrapping the DNA around protein particles known as nucleosomes. Most DNA molecules are linear but some DNAs are circles, as is often the case for the chromosomes of prokaryotes and for certain viruses.

DNA is flexible. Unless the molecule is topologically constrained, it can freely rotate to accommodate changes in the number of times the two strands twist about each other. DNA is topologically constrained when it is in the form of a covalently closed circle, or when it is entrained in chromatin. The linking number is an invariant topological property of covalently closed circular DNA. It is the number of times one strand would have to be passed through the other strand in order to separate the two circular strands. The linking number is the sum of two interconvertible geometric properties: twist, which is the number of times the two strands are wrapped around each other; and the writhing number, which is the number of times the long axis of the DNA crosses over itself in space. DNA is relaxed under physiological conditions when it has about 10.5 base pairs per turn and is free of writhe. If the linking number is decreased, then the DNA becomes torsionally stressed, and it is said to be negatively supercoiled. DNA in cells is usually negatively supercoiled by about 6%.

The left-handed wrapping of DNA around nucleosomes introduces negative supercoiling in eukaryotes. In prokaryotes, which lack histones, the enzyme DNA gyrase is responsible for generating negative supercoils. DNA gyrase is a member of the type II family of topoisomerases. These enzymes change the linking number of DNA in steps of two by making a transient break in the double helix and passing a region of duplex DNA through the break. Some type II topoisomerases relax supercoiled DNA, whereas DNA gyrase generates negative supercoils. Type I topoisomerases also relax supercoiled DNAs but do so in steps of one in which one DNA strand is passed through a transient nick in the other strand.

RNA differs from DNA in the following ways: its backbone contains ribose rather than 2′-deoxyribose; it contains the pyrimidine uracil in place of thymine; and it usually exists as a single polynucleotide chain, without a complementary chain. As a consequence of being a single strand, RNA can fold back on itself to form short stretches of double helix between regions that are complementary to each other. RNA allows a greater range of base pairing than does DNA. Thus, as well as A:U and C:G pairing, U can also pair with G. This capacity to form a non-Watson-Crick base pair adds to the propensity of RNA to form double-helical segments. Freed of the constraint of forming long-range regular helices, RNA can form complex tertiary structures, which are often based on unconventional interactions between bases and between bases and the sugar-phosphate backbone.

Some RNAs act as enzymes—they catalyze chemical reactions in the cell and in vitro. These RNA enzymes are known as ribozymes. Most ribozymes act on phosphorous centers, as in the case of the ribonuclease RNase P. RNase P is composed of protein and RNA, but it is the RNA moiety that is the catalyst. The hammerhead is a self-cleaving RNA, which cuts the RNA backbone via the formation of a 2′, 3′ cyclic phosphate in a reaction that involves an RNA-bound Mg^{++} ion. Peptidyl transferase is an example of a ribozyme that acts on a carbon center. This ribozyme, which is responsible for the formation of the peptide bond, is one of the RNA components of the ribosome. The discovery of RNA enzymes that can act on phosphorous or carbon centers suggests that life might have evolved from a primitive form in which RNA functioned both as the genetic material and as the enzymatic machinery.

PROBLEMS

1. Draw an A:T base pair.

a. Indicate how each base is joined to deoxyribose.

b. Indicate which edge of each base pair faces into the major groove and which into the minor groove.

c. Use two different colors to indicate whether an atom is a hydrogen bond donor or a hydrogen bond acceptor.

2. Draw a G:C base pair.

a. Indicate how each base is joined to deoxyribose.

b. Indicate which edge of each base pair faces into the major groove and which into the minor groove.

c. Use two different colors to indicate whether an atom is a hydrogen bond donor or a hydrogen bond acceptor.

3. Other than hydrogen bonding, what else contributes to the stability of the double helix?

4. Certain chemical agents such as nitrous acid can deaminate cytosine, converting it into uracil. How might this explain why DNA contains thymine in place of uracil?

5. The virion DNA of an *E. coli* phage called ϕX174 has the base composition: 25% A, 33% T, 24% G, and 18% C. What do these data suggest about the structure of the phage's chromosome?

6. Describe several reasons why the major groove is more often used by proteins to recognize specific DNA sequences than the minor groove. Consider the sequence AATCGG; what information, in terms of hydrogen bond donors, hydrogen bond acceptors, nonpolar hydrogen, and methyl groups, are provided by the major groove and minor groove, in each direction?

7. Describe the ways in which DNA can vary from its ideal B structure, and contrast the B form of DNA with the A and Z forms. What factors or conditions favor these deviations from the B form and the choice between the three possible forms?

8. Draw a graph showing the OD_{260} as a function of temperature for DNA isolated from a bacterial species having a high GC content, and one from a bacterial species having a low GC content.

9. In a collaborative project with a physicist colleague of yours, you decide to test the effect of a new element he has discovered, fictionium, on the pitch of DNA in solution. You are familiar with the mica experiment described in the text, in which DNA is bound to mica and the exposed side is cut with DNase I. You attempt to measure the pitch of your DNA using the same method, but much to your dismay you realize that fictionium strongly inhibits DNase I, and so your experiment is an utter failure. Before you report the bad news to your colleague, what alternative experimental approaches could you take to determine this value?

10. Consider a covalently closed, circular DNA molecule of length 10,500 base pairs and *Lk* 950. What is the effect of the binding of 110 molecules of ethidium on *Lk*, *Tw*, and *Wr*?

11. Which of the following structures have twist, which have writhe, and which have both?

a. a closed circular DNA molecule lying flat on a plane

b. double-stranded DNA wrapped around a nucleosome

c. a circular, single-stranded oligonucleotide

d. an overtwisted telephone cord

e. a human chromosome

12. Describe three differences between topoisomerase I and topoisomerase II. You have an experiment in mind that requires topoisomerase II, but not topoisomerase I, and would like to purify this enzyme from bacterial cells. Describe a purification strategy that would allow you to specifically isolate topoisomerase II, relying on the unique activities of each enzyme.

13. Populations of the following types of molecules are incubated with the indicated enzymes. Predict all possible products for each reaction.

a. Complementary single-stranded circles + Topoisomerase I

b. Negatively supercoiled DNA + eukaryotic Topoisomerase II

c. Negatively supercoiled DNA + Topoisomerase I

14. Draw the reaction that causes RNA to hydrolyze at high pH.

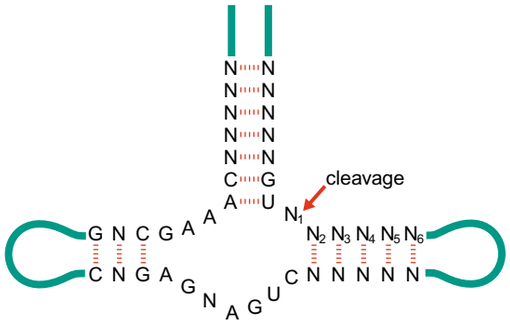
a. Why is RNA more sensitive to high pH than DNA?

b. What is the function of Mg^{2+} in RNA molecules?

15. While RNAse P contains both RNA and protein, the enzymatic activity is known to reside in the RNA component. What is the role of the protein, then, in this enzyme? What kind of experiment could be used to demonstrate that the activity of the enzyme resides in the RNA component, and not in the protein?

16. The so-called hammerhead ribozyme mediates its own cleavage. It was initially identified in plant virions and has been shown to have the secondary structure depicted in the figure below.

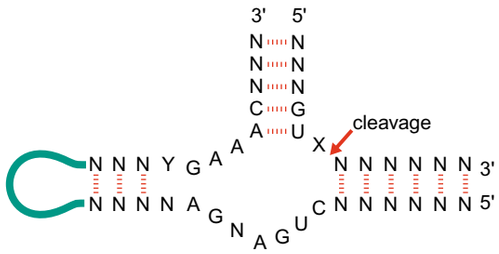
The position of cleavage is marked by the arrow in the figure. (Note that N_i can be any base, and the subscript *i* is used to denote N at a particular position.) Describe the chemical nature of the cleavage reaction, indicating which chemical group of which base is attacking which other chemical group.



17. Instead of using a single-stranded molecule like that shown in Problem 16, a two-stranded hammerhead, as shown in the figure below, is often used.

a. What is the major consequence of using a two-stranded structure?

b. Is it feasible for a single-stranded DNA to be an enzyme? Explain. How might the lack of a 2′-OH group in DNA be remedied?



18. The Fragile-X syndrome is the most common inherited form of mental retardation in humans. The gene causing the disease has been cloned and shown to encode an RNA-binding protein that binds to a diverse yet specific pool of mRNA species in the brain. Based on what you know about RNA structure, do you think it likely that this protein binds to these RNA molecules using a similar mechanism that proteins use to bind DNA? Explain why or why not, and if not, propose another way that this protein may recognize these RNA species.

19. Recall the approach described in Box 6-2 that allowed us to conclude that DNA in solution has a helical periodicity of 10.5 base pairs per turn of the helix. Now consider scenarios in which the solution to the question of helical periodicity is either 10 or 11 base pairs, rather than 10.5. Design experiments comparable to those described in Box 6-2, using cccDNAs of the appropriate lengths, and provide anticipated experimental data which would suggest that DNA has a helical periodicity of 10 base pairs or of 11 base pairs per turn of the helix.

BIBLIOGRAPHY

Books

Kornberg, A. and Baker, T. A. 1992. *DNA Replication*. New York: W.H. Freeman.

Lodish, H. et al. 1999. *Molecular Cell Biology*, 4th edition. New York: W. H. Freeman and Company.

Gesteland, R. F. et al., ed. 1992. *The RNA World*, 2nd edition. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press.

Saenger, W. 1984. *Principles of Nucleic Acid Structure*. New York: Springer-Verlag.

Sarma, R. H., ed. 1981. *Bimolecular Stereodynamics*, Vols. 1 and 2. Guilderland, N.Y.: Adenine Press.

Cold Spring Harbor Symposium on Quantitative Biology. 1982. Volume 47: Structures of DNA. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press.

DNA Structure

Dickerson, R. E. 1983. The DNA helix and how it is read. *Sci. Amer.* **249**:94–111.

Franklin, R. E. and Gosling, R. G. 1953. Molecular configuration in sodium thymonucleate. *Nature* **171**:740–741.

Rich, A., Nordheim, A., and Wang, A. H. J. 1984. The chemistry and biology of left-handed Z DNA. *Annu. Rev. Biochem*. **53**:791–846.

Roberts, R. J. 1995. On base flipping. *Cell* **82(1)**:9–12.

Wang, A. H., Fujii, S., van Boom, J. H., and Rich, A. 1983. Right-handed and left-handed double-helical DNA: Structural studies. *Cold Spring Harb. Symp. Quant. Biol*. **47** Pt 1: 33–44.

Watson, J. D. and Crick, F. H. C. 1953. Molecular structure of nucleic acids: A structure for deoxyribonucleic acids. *Nature* **171**:737–738.

———. 1953. Genetical implications of the structure of deoxyribonucleic acids. *Nature* **171**:964–967.

Wilkins, M. H. F., Stokes, A. R., and Wilson, H. R. 1953. Molecular structure of deoxypentose nucleic acids. *Nature* **171**:738–740.

DNA Topology

Bauer, W. R., Crick, F. H. C., and White, J. H. 1980. Supercoiled DNA. *Sci Amer.* **243**:118–133.

Boles, T. C., White, J. H., and Cozzarelli, N. R. 1990. Structure of plectonemically supercoiled DNA. *J. Mol. Biol.* **213**:931–951.

Crick, F. H. C. 1976. Linking numbers and nucleosomes. *Proc. Natl. Acad. Sci.* **73**:2639–2643.

Dröge, P. and Cozzarelli, N. R. 1992. Topological structure of DNA knots and catenanes. *Methods Enzymol.* **212**: 120–130.

Gellert, G. H. 1981. DNA topoisomerases. *Annu. Rev. Biochem*. **50**:879–910.

Wang, J. C. 2002. Cellular roles of DNA topoisomerases: A molecular perspective. *Nat. Rev. Mol. Cell Biol*. **3**:430–440.

Wasserman, S. A. and Cozzarelli, N. R. 1986. Biochemical topology: Applications to DNA recombination and replication. *Science* **232**:951–960.

RNA Structure

Doherty, E. A. and Doudna, J. A. 2001. Ribozyme structures and mechanisms. *Ann. Rev. Biophys. Biomol. Struct.* **30**:457–475.

McKay, D. B. and Wedekind, J. E. 1999. Small ribozymes. In *The RNA World*, 2nd edition (ed. Gesteland, R. F. et al.), pp. 265–286. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press.

Uhlenbeck, O. C., Pardi, A., and Feigon, J. 1997. RNA structure comes of age. *Cell* **90**:833–840.