

DNA Structures and Spectra

The study of the physico-chemical bases of the interaction of drugs with DNA, as well as the expression of genes and the genetic inheritance of biological characteristics, began in earnest with the quest for the structure of DNA itself, which was ultimately “solved” in 1953. The earliest research on DNA structure used optical spectroscopic methods, which are still used for studying drug–DNA interactions. By 1950, Maurice Wilkins and Ray Gosling began to study DNA structure using X-ray diffraction and obtained the first diffraction images of DNA fibers. With help from Alec Stokes, their colleague at Kings College, they proposed that the patterns resulted from a helical structure.

In 1951, Rosalind Franklin also joined the faculty at Kings College. Franklin worked closely with Wilkins and Gosling, but by using her own techniques, she obtained better X-ray images of DNA than anyone had done previously. These images clearly showed the helical structure of DNA, and enabled the identification of the locations of the phosphate sugars. These data also revealed that the DNA had two forms, depending on environmental humidity, which Franklin and Wilkins called the A and B forms of DNA. Meanwhile, Francis Crick and James Watson of Cambridge University were also working on DNA structure. They obtained some of Franklin’s data and, by combining her data with their own analyses, built the first correct DNA model. In 1953, their paper was published in *Nature* [1], along with papers by Wilkins, Stokes, and Wilson [2], and Franklin and Gosling [3]. Sadly, Franklin succumbed to cancer in 1958, but Wilkins, Watson, and Crick shared the Nobel Prize in Physiology and Medicine in 1962.

It is likely that modern molecular genetics would not exist without this basic structural knowledge, and since that time, spectroscopic studies have continued to expand our physical understanding of DNA structure. There remains growing need to understand the structure of DNA and its interactions with drugs, mutagens, RNA, and protein ligands. DNA is no longer regarded as a rigid cylinder, but as a highly plastic and adaptable molecule, capable of many structural variations, each having unique and important biological consequences.

The aim of this chapter is to explain the most fundamental concepts of DNA structure, especially the terminology and factors that describe its primary, secondary, tertiary, and quaternary structures. It will describe the structures and spectroscopic properties of DNA, as well as relate these to its biophysical characteristics.

1.1. DNA STRUCTURES

1.1.1. Primary (Chemical) Structure

The monomeric units of nucleic acids are called *nucleotides*. These consist of a phosphate, a sugar, and a special aromatic moiety called a purine (*pur*) or pyrimidine (*pyr*) base. Figure 1.1 shows the chemical structures and names of the common bases and sugars and also the numbering of the ring atoms (C and N) of bases. This IUPAC [4] numbering system is used throughout this book. All of the carbon atoms in the DNA bases and sugars are given unique position numbers. This is also true of the nitrogen atoms in the aromatic rings of the bases. All other atoms or groups (H, O, NH₂, and P) are numbered according to the carbon or nitrogen atoms to which they are directly bound. The thymine methyl group is given the unique number of C7 by IUPAC. (An alternate number of C5M is used in the PDB database, but will not be used in this book.) The sugars have the β -D-stereoisomeric configuration shown in Figure 1.1, which is a reference to the relative orientations of the OH groups. The atoms belonging to the

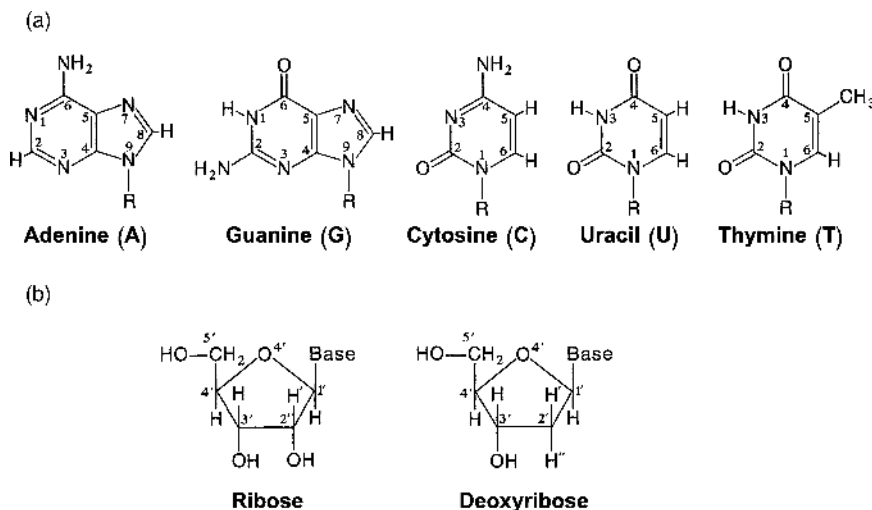


FIGURE 1.1. Chemical structure of nucleic acid constituents, the normal bases (a) and sugars (b). In DNA and RNA, the R group is the attached sugar ring. Isolated bases have R=H. The IUPAC [4] nomenclature for naming atoms are given here, and is used exclusively in this book. All C and aromatic N have unique atom position numbers, whereas the other atoms and groups (H, O, NH₂, and P) are numbered according to the carbon or nitrogen atoms to which they are directly bound. The thymine methyl group is given the unique number of C7 by IUPAC.

sugar moiety are distinguished from those of the base by one or two superscript prime marks (i.e., ' or ") on the atom number. The chemical unit that contains only the sugar and base is called a *nucleoside*; the phosphate ester of a nucleoside is called a *nucleotide*.

Nucleosides are abbreviated by combining an uppercase letter representing the base and a lowercase letter representing the sugar. For example, adenosine, which has a ribose sugar and an adenine base, is called "rA" (see Fig. 1.2). Deoxycytidine, with a 2'-deoxyribose sugar and a cytosine base, is "dC" (Fig. 1.2). Nucleotides are denoted by adding a "p" to represent each phosphate, as shown in Fig. 1.2; thus, for 5'-nucleotide, the abbreviation is "pN," and for 3'-nucleotide, "Np." If the base in Fig. 1.2 (bottom left) is guanine, it is referred to as guanosine-5'-monophosphate, often abbreviated as 5'-GMP.

Natural nucleic acids, RNA and DNA, are polynucleotides, the backbone of which consists of alternating sugars and phosphodiester, as shown in Fig. 1.3. In RNA (ribonucleic acid) the sugars are ribose, and four nucleosides, rA, rU, rC, and rG, predominate. In DNA (deoxyribonucleic acid), the sugars are all 2'-deoxyribose, and four nucleosides, dA, dT, dC, and dG, make up essentially the total structure of all known DNAs. For both of these polynucleotides, each sugar has two ester linkages at the 5'O and 3'O atoms. The DNA chain is described as having a "direction," which is defined as the contiguous set of atoms starting at the 5'O and going to the 3'O ($5' \rightarrow 3'$).

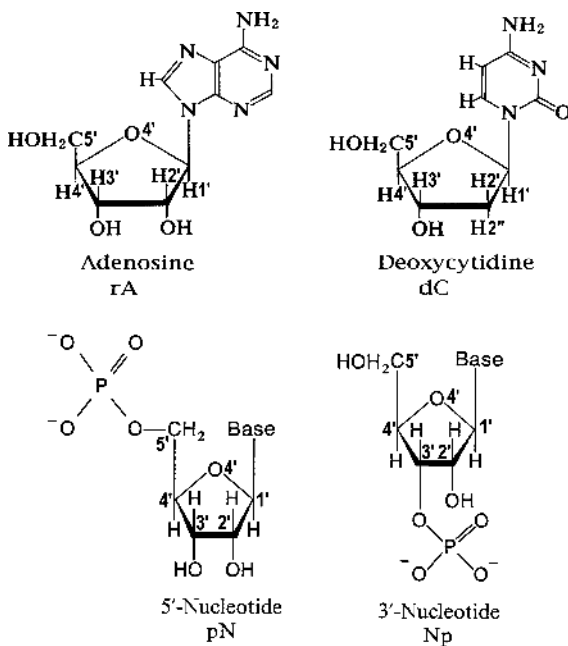


FIGURE 1.2. Chemical structure of nucleoside and nucleotide. A common nucleoside analog is inosine, which is similar to similar to guanosine, except that the —NH_2 group is replaced by an H.

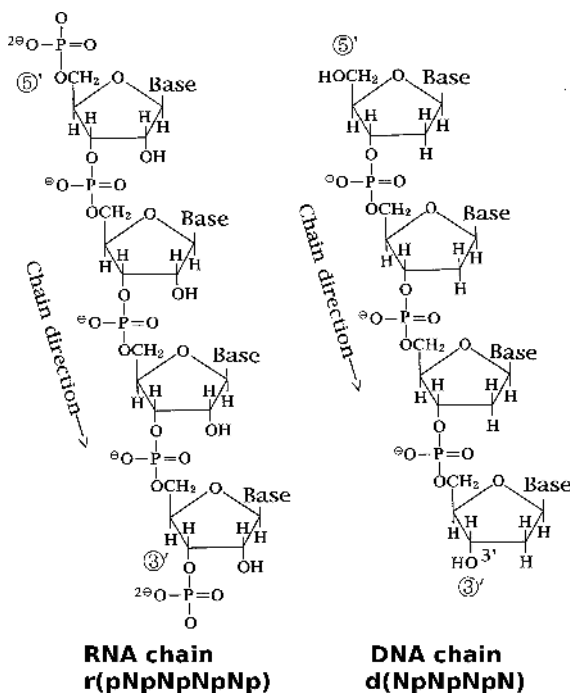


FIGURE 1.3. Chemical structure of polynucleotide. The terminal oxygen atoms, O5' and O3', are named according to the carbon atoms to which they are directly bound.

(Note that these oxygen atoms are designated as 5' and 3', according to the carbon atom of the sugar to which they are bound.) This set of atoms is called the "backbone" of the DNA. Thus, a polynucleotide is abbreviated as ...pNpNpNpNpNpNp..., where the left end is called the "5' end" and the right end, the "3' end." These are also known as "the five-prime end" and "the three-prime end."

All the common DNA bases are uncharged at neutral pH. At lower pH, however, adenine (A) and cytosine (C) have positively charged protonated NH^+ forms, as shown in Fig. 1.4. Uracil (U) and thymine (T) have negatively charged deprotonated forms (N^-) at higher pH, while guanine (G) has both a positively charged protonated form (at lower pH) and a negatively charged deprotonated form (at higher pH; see Fig. 1.4).

In contrast, all the phosphate residues in nucleotides or polynucleotides are negatively charged at the pH at which one normally works with these compounds. This means that there must always be sufficient counter-cations present in the environment to neutralize the electronic charges on the DNA. The monoester and diester forms of phosphoric acid exist in an ionization equilibrium, as shown in Fig. 1.5. Therefore, in every nucleotide, at either end of a polynucleotide (5' or 3'), the phosphate has a doubly negatively charged form (DNA-O-PO_3^{2-}) at pH higher than 7 (i.e., at $\text{pH} > 7$). On the other hand, all of the phosphodiester linkages in the polynucleotide chain have a single negative charge ($-\text{O-PO}_2^--\text{O}-$) at neutral pH.

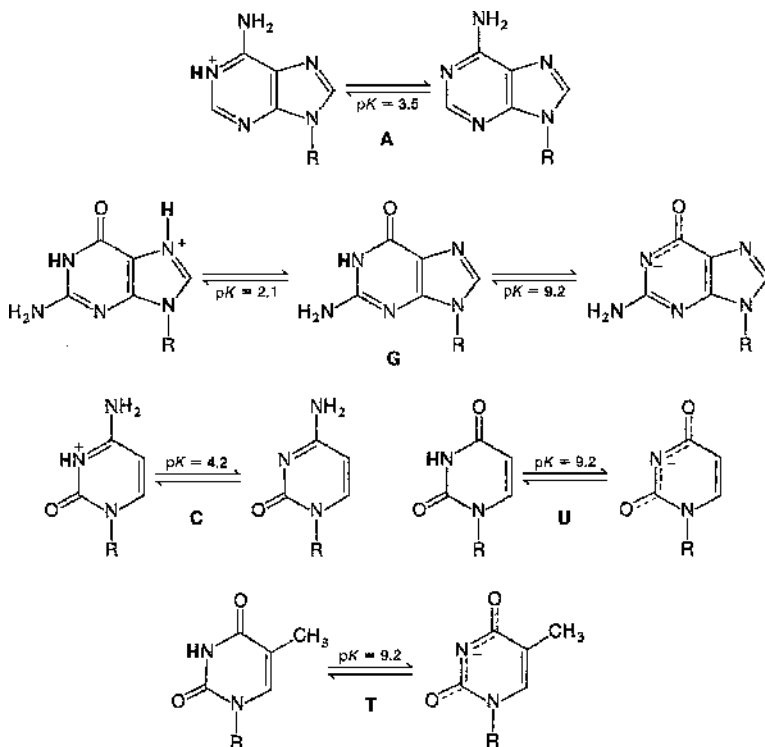


FIGURE 1.4. Ionization of aromatic bases (nucleosides) in DNA and RNA.

It should be noted that the two nonester P—O bonds are nearly equivalent to each other, so that each of these bonds has semi-double-bond character, and the single negative charge is shared between the two oxygen atoms.

1.1.2. Bond Lengths and Bond Angles

The secondary structure of DNA (or any polynucleotide) consists of three-dimensional regions that have ordered, locally symmetric, structures. The local

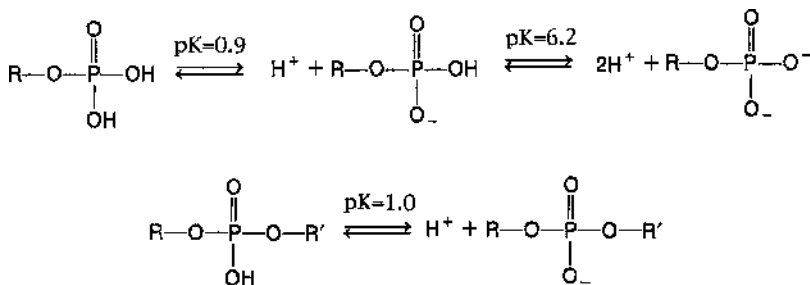


FIGURE 1.5. Ionization of phosphate.

three-dimensional structure is given by a list of atomic coordinates (see discussion below and Table 1.2, at the end of Section 1.1.3).

The three-dimensional structure of a DNA (or polynucleotide) molecule can be fully defined by listing either the (1) atomic Cartesian coordinates, or (2) the bond lengths, the bond angles, and the internal dihedral torsion angles. While the former type of list (1) is simpler to write out, the latter (2) better clarifies the physico-chemical properties, and has the advantage of listing characteristics that are not limited to a specific location in space, but that can be freely translated or rotated. Dihedral torsion angles are explained more fully in Section 1.1.8. The term *dihedral torsion angles* is often shortened in the literature, to either *torsion angles* or *dihedral angles*. In this book, we will use the term *dihedral angles* throughout.

All nucleic acid base residues are nearly planar, having deviations from planarity of less than 0.2 Å. Typical bond lengths and angles for the nucleoside base residues are given in Fig. 1.6 [5]. These aromatic moieties are central to the genetic significance of DNA, and contribute significantly to its energetic stability through the inter-residue van der Waals attractive forces between the π electrons in the aromatic rings. “Stacking interactions” of this type are revisited throughout this book.

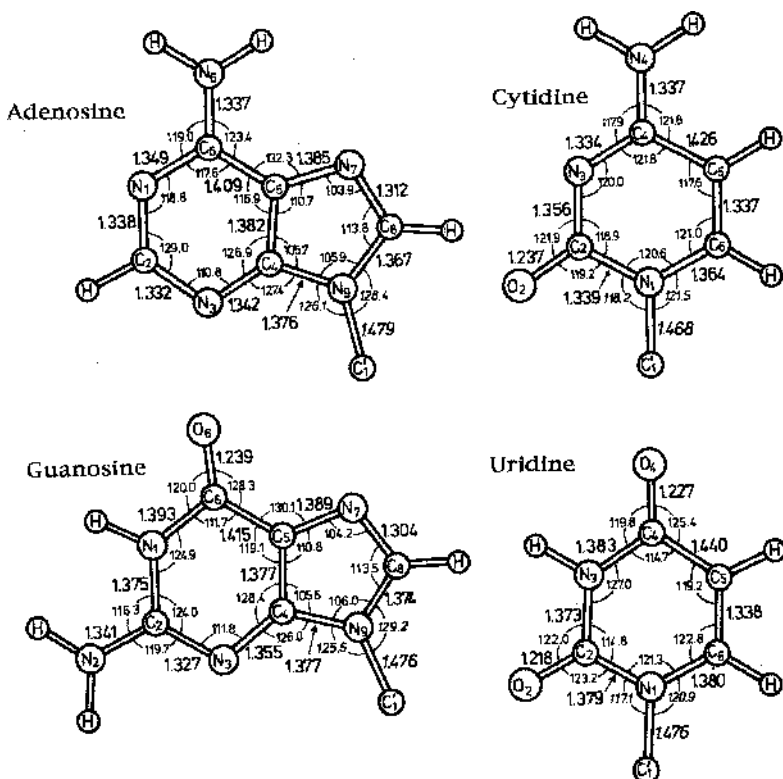


FIGURE 1.6. Average bond lengths (Å) and bond angles (°) in 1-substituted uracil and cytosine and in 9-substituted adenine and guanine [10].

Unlike the nucleotide bases, the furanose sugar rings are not planar, but have out-of-plane puckering in either an “envelope” or a “chair” conformation (Section 1.1.4). To better understand this puckering, it is instructive to examine the conformation of tetrahydrofuran (Fig. 1.7a), which is the unsubstituted prototype of the five-membered sugar rings in nucleosides. According to electron diffraction [6,7], microwave [8], and theoretical [9] studies, this simple five-membered ring is also not planar, but is puckered. The bond distances agree well with standard, alkane-like C–C and C–O values (see Fig. 1.7), but it is the bond angles that cause the nonplanarity. These angles are

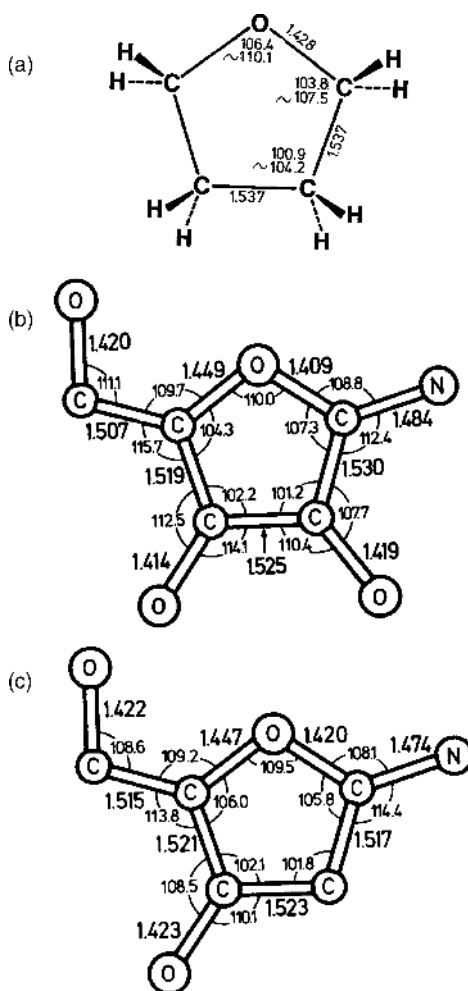


FIGURE 1.7. (a) Geometry of the unsubstituted furanose five-membered ring, tetrahydrofuran, as obtained from electron diffraction data [6–9]; (b) geometric data for ribose in nucleosides with C3'-endo pucker; (c) geometric data for deoxyribose in nucleosides with C2'-endo pucker [10].

significantly smaller than both the standard sp^3 -tetrahedral geometry (109.5°) and the expected internal angles for a planar five-membered ring ($540^\circ/5 = 108^\circ$). In addition, the bond angles are not fixed, but fluctuate in the ranges shown in Fig. 1.7a. Much of the flexibility of the DNA backbone is attributable, at least in part, to these sugars.

The C—C and C—O bond distances in nucleoside furanose sugars are nearly the same as those in tetrahydrofuran, whereas more pronounced differences are seen in the bond angles (Fig. 1.7). The five-membered sugar ring can acquire a variety of puckering structures, which will be described in detail in Section 1.1.4, following discussion of the canonical three-dimensional structure of DNA in Section 1.1.3.

The phosphate groups bridging between nucleosides are phosphodiester, in which the ester (P—O) “single” bonds have some double-bond character due to π -electron delocalization. This is reflected in the fact that the P—O bonds in these mono- and diesters range within 1.59–1.63 Å (Fig. 1.8) [5,10], which is an interatomic distance

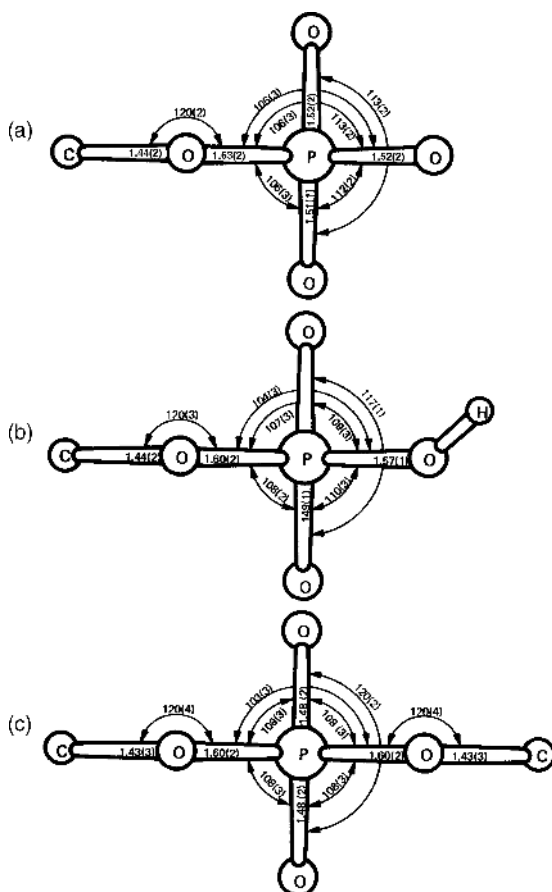


FIGURE 1.8. Geometric data for phosphoric acid esters: (a) monoester in higher pH; (b) monoester in lower pH (protonated); (c) diester [10].

that is slightly shorter than that of standard single bonds. Likewise, the valence P—O—C bond angle is $\sim 120^\circ$, which is wider than expected for an oxygen atom with sp^3 character. Since these are single bonds with some double-bond characteristics, they have the effect of allowing a somewhat reduced amount of torsional flexibility in this portion of the DNA backbone.

In the single protonated phosphomonoesters (Fig. 1.8b), which occur at the very ends of the DNA chain, the P—OH bond distance is close to 1.57 Å. This is shorter than the P—O ester bond but longer than the two P—O bonds that carry the negative charge. The two unesterified P—O bond distances (1.46–1.56 Å) vary as a result of environmental influences in the crystalline state and are distinctively longer than the P=O double bond of 1.40 Å. If the phosphomonoester is not protonated but exists as a dianion (Fig. 1.8a), then the three unesterified P—O bonds are electronically equivalent. Likewise, the two unesterified P—O bonds in the diester anion (Fig. 1.8c), which links the DNA residues, are also electronically equivalent.

1.1.3. Canonical Structure of DNA: B-DNA

DNA has significant structural plasticity and can adopt a wide range of three-dimensional conformations. DNA rarely exists as a single strand, but prefers to associate into double-stranded, or “duplex” structures, the most common of which are known as the “A,” “B,” and “Z” conformations. More than 90% of known DNA sequences have three-dimensional conformations that fall into a family of structures called “B-DNA.” For example, the cellular DNA found in *Escherichia coli*, virus P22, and the chromatin of all higher organisms are all B-form structures. As described in Section 1.1, the polynucleotide chain is a heavily charged polyanion that is neutralized by cations such as K^+ and Na^+ . These cations can have a significant effect on DNA structure, but as a first approximation, one usually considers the generalized B-DNA structure to be independent of the kind of counter-ion.

The polynucleotide chain of B-form DNA is characterized by a “right-handed” helix. This term derives from the fact that when one points along the central, helical axis with one’s right thumb, the fingers curl in the direction of the twist of the strand. For a right-handed helix, the strands curl counterclockwise when the helical axis is pointing upward. In B-form DNA, two such helices run in antiparallel directions with respect to each other, and are interwound coaxially (Fig. 1.9b). The 5′ end of one strand is paired with the 3′ end of the other, in a head-to-tail orientation. The two helices are held together through hydrogen bonds between two base residues, one from each strand, forming specific base pairs. The B-form helix makes one complete turn in about 10 residues, has a diameter of about 24 Å, and has a “pitch” of about 35.5 Å (Table 1.1). The pitch of a helix is the amount of rise necessary for one complete turn of 360° .

The standard base pairs consist of guanine to cytosine (G-C) and adenine to thymine (A-T). These are called *Watson–Crick base pairings* (Fig. 1.10) (for other possible pairings, see Section 1.1.11). These bases are aromatic moieties with internal π -electron delocalization. This delocalization both flattens the individual bases and stabilizes the vertical stacking of these bases. Such stacking interactions are an

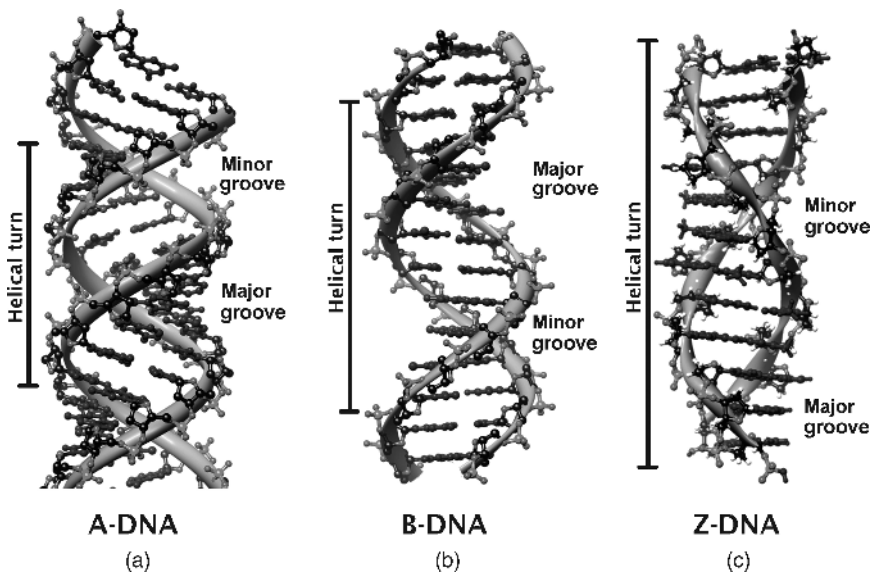


FIGURE 1.9. The side views of idealized fragments of (a) A-DNA, (b) B-DNA, and (c) Z-DNA, along with their approximate helical turns [276]. The A form has the tightest helical turn and hence is the shortest. Its base pairs also have the greatest amount of tilt with respect to the helical axis. The helical turn for the Z form is the least compact and stretches the longest distance. Note that the alternation of *syn* and *anti* glycosidic dihedral angles within the Z form is emphasized here as a twisting of the backbone ribbon. The B form has a moderate helical turn and length, with regular glycosidic dihedral torsion angles. (See color insert.)

TABLE 1.1. Geometric Characteristics of Different DNA Helical Forms

Structural Characteristic	Form		
	A	B	Z
Helical sense	Right-handed	Right-handed	Left-handed
Diameter	25.5 Å	23.7 Å	18.4 Å
Repeating unit	1 BP	1 BP	2 BP
Rotation per repeating unit (t_g)	32.7°	35.9°	−60°
Base pairs (BP) per turn (mean)	10.7	10.5	12
Tilt of BP relative to helical axis	+19°	−1°	−9°
Rise/BP along axis (p)	2.56 Å	3.38 Å	3.71 Å
Pitch/turn of helix (P)	28.2 Å	35.5 Å	44.6 Å
Mean propeller twist	+18°	+16°	0°
Sugar pucker	C3'- <i>endo</i>	C2'- <i>endo</i>	G: C3'- <i>endo</i> C, T, A: C2'- <i>endo</i>
Glycosidic dihedral torsion angle (χ)	<i>anti</i>	<i>anti</i>	G: <i>syn</i> C, T, A: <i>anti</i>
Minor groove	Wide and shallow	Narrow and deep	Narrow and deep
Major groove	Deep and narrow	Deep and wide	Shallow

Source: Data compiled from Refs. 10, 67, 198 and 277.

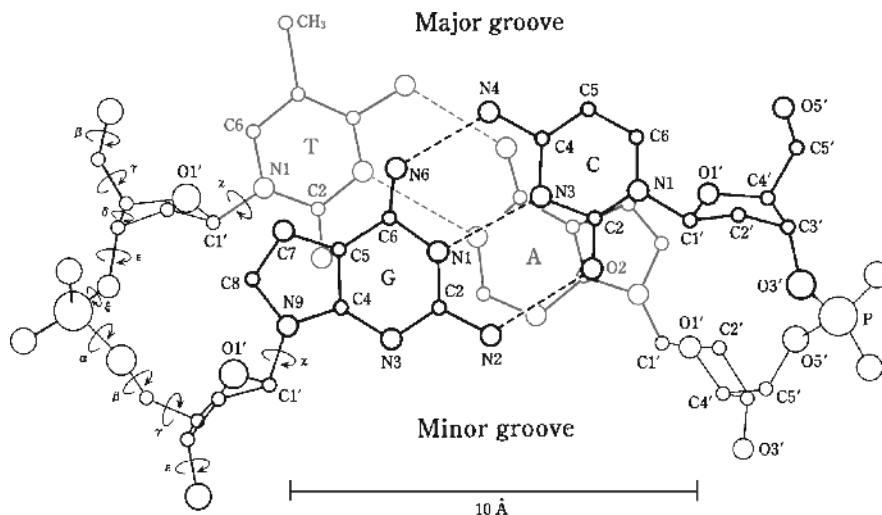


FIGURE 1.10. B-DNA structure, viewed along the helical axis. A G-C base pair is “above” an A-T base pair. (See color insert.)

important source of the energetic stability of DNA. The base pairs form planes that are stacked on top of one another along the axis of the double helix, and are oriented roughly perpendicular to the helical axis. As will be discussed later, the base pairs rarely form perfect planes; rather, the two bases are usually slightly twisted with respect to each other in a propeller-like manner. Not surprisingly, the angle between the planes of the two bases is a measure of a phenomenon called the “propeller twist.”

As noted above, the phosphate groups of DNA are negatively charged, and as there are many phosphate groups in the DNA strands, the overall molecule has a very large net charge. The electrostatic repulsion between two DNA strands is sufficient to prevent their association into a duplex. To counteract this tendency, a sufficient concentration of cations must be present to maintain electroneutrality. The primary cations in the cellular environment are the monovalent ions, K^+ and Na^+ , which are hydrated in solution. They stabilize DNA by electrostatically clustering around the phosphate groups, forming a loose spine along the strand. This clustering is referred to *salt condensation*. Divalent cations, such as Mg^{2+} , can also stabilize DNA, but often do so by bridging between the two phosphate groups on opposite strands. For this reason, divalent cations can provide greater stability to DNA, sometimes stabilizing nonstandard DNA structures, such as distorted or triple helices.

Many of the most important structural characteristics of B-form DNA, as well as those of A- and Z-form DNA (see later sections), are summarized in Table 1.1. The atomic coordinates in B-DNA were given first by Wilkins and his collaborators [11–13], and later revised by Arnott and others [14–16]. Table 1.2 lists the atomic coordinates of B-DNA given by Arnott and Hukins [15]. These are the geometrically preferred structural parameters (bond lengths, bond angles, dihedral angles, etc.) describing a DNA structure that is linear and not distorted by drug binding.

TABLE 1.2. Cylindrical Polar Coordinates of a Nucleotide Residue in B-DNA^a

Atom		r (Å)	θ (°)	z (Å)
Phosphate	O1	8.75	97.4	3.63
	O2	10.20	91.1	1.86
	O3	8.82	103.3	1.29
	P1	8.91	95.2	2.08
Deoxyribose	O4	7.73	88.0	1.83
	C5	7.70	79.8	2.77
	O5	6.22	66.0	1.83
	C4	7.59	69.9	2.04
	C3	8.20	69.9	0.64
	C2	7.04	73.2	−0.24
	C1	5.86	67.4	0.47
	N9	4.63	76.6	0.42
Adenine	C8	4.84	93.0	0.50
	N7	3.95	105.4	0.43
	C5	2.74	94.0	0.28
	N6	1.83	154.0	0.14
	C6	1.41	107.2	0.15
	N1	0.86	40.1	0.03
	C2	2.17	30.6	0.04
	N3	3.24	47.0	0.16
	C4	3.33	70.5	0.28
	N9	4.63	76.6	0.42
Guanine	C8	4.82	93.2	0.50
	N7	3.92	105.7	0.42
	C5	2.70	94.0	0.28
	O6	1.71	154.6	0.13
	C6	1.39	109.3	0.15
	N1	0.92	37.9	0.03
	N2	3.01	4.2	−0.10
	C2	2.28	28.7	0.03
	N3	3.29	46.7	0.16
	C4	3.33	70.3	0.28
Cytosine	N1	4.63	76.6	0.42
	C6	4.99	92.2	0.52
	C5	4.35	107.0	0.47
	N4	2.76	136.6	0.27
	C4	2.94	110.0	0.32
	N3	2.31	83.9	0.22
	O2	3.69	47.9	0.18
	C2	3.40	67.4	0.27
Thymine	N1	4.63	76.6	0.42
	C6	5.01	92.3	0.52
	Me	5.40	119.8	0.58
	C5	4.38	106.9	0.47
	O4	2.82	136.3	0.27
	C4	2.98	111.9	0.32
	N3	2.36	85.2	0.23

TABLE 1.2. (Continued)

Atom	r (Å)	θ (°)	z (Å)
O2	3.64	47.8	0.18
C2	3.42	67.3	0.27

^aCoordinates for successive nucleotide residues may be obtained by adding 36.0° to θ and 3.38 Å to z . All base coordinates refer to one residue. To generate coordinates of the dyadically related base pair, change the signs of θ and z .
Source: Ref. 15.

1.1.4. Deoxyribose Moiety as a Flexible Joint

To be biologically relevant, a DNA duplex must maintain a sufficiently regular structure to be recognized by proteins as a line of letters spelling genetic information. At the same time, however, it must also have enough flexibility to allow its strands to dissociate and associate during transcription and replication. One source of this flexibility is a peculiarity of the deoxyribose moiety. The sugars act as a junction between the phosphate groups and the sidechain (the base residue) and can function as a partial “universal joint,” which provides just the right amount of flexibility to the whole DNA structure.

As mentioned in Section 1.1.2, a saturated, five-membered ring (like deoxyribose ring) is not planar, but has an envelope (E) or a twist (T) geometric conformation (Fig. 1.11). In an E conformation, one of the member atoms (C or O) protrudes from the plane formed by the other four atoms; in a T conformation, two adjacent members are out of plane, with one above the plane (up) and the other below it (down). Ten possible E forms of the deoxyribose ring are shown in Fig. 1.12: C1'-*endo*, C2'-*exo*, and so on. Here, “C1'-*endo*” refers to a geometry in which the C1' atom is projected outward and upward (toward the same side on which the glycosidic bond is located) with respect to the C2'C3'C4'O plane. It also is denoted as ¹E, where the “E” refers to the envelope form and “1” refers to C1'. The fact that the “1” is raised (superscripted), reminds us that the C1' is oriented *upward*, or in the *endo* conformation. The term “C2'-*exo*” refers to a geometry in which the C2' atom is projected downward (toward the side opposite to the location of the glycosidic bond) from the C1'C3'C4'O plane. It has

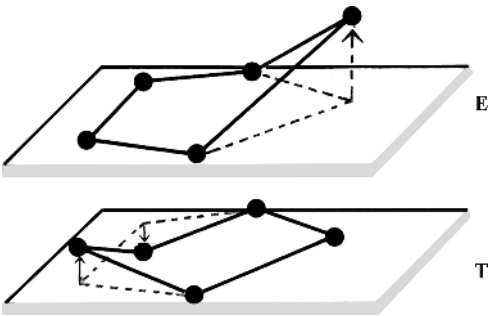


FIGURE 1.11. Two possible conformations of a saturated five-membered ring (or a deoxyribose ring) (E—envelope form, T—twist form).

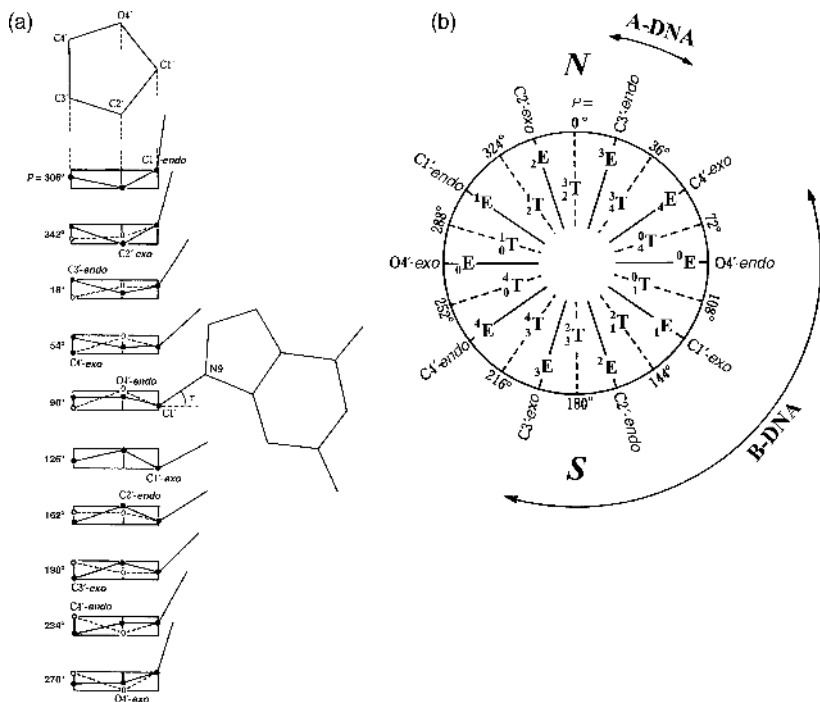


FIGURE 1.12. Conformations of the deoxyribose ring. (a) Side views of each of the 10 defined envelope forms, with the associated P (= angle of pseudorotation) value. Carbon atoms and bonds are indicated by solid lines and dots (●) if they are on the reader's side of the ring, while dashed lines and hollow dots (○) indicate bonds and atoms located on the far side of the ring. An "O" indicates an oxygen atom located in the far side of the ring. (b) The relation between the pseudorotation angle P and the conformation of deoxyribose ring. E is an envelope conformation. N indicates "north," or $P \sim 0$, and S indicates "south," or $P \sim 180^\circ$.

the label of ${}_2E$, indicating that it is an envelope form with the C_2' oriented *downward*, or in the *exo* conformation.

Since the internal dihedral angles within the deoxyribose ring are all interdependent, a change in one of the angles must affect all the others. Only two parameters are needed to define the conformations of all the dihedral angles (z_i) within the ring. The first parameter ($z_{\max}$), defines the maximum out-of-plane pucker, or maximum dihedral angle within the ring. The other parameter is the pseudorotation phase angle, or P . Every E or T form is regarded as a "stop" in the *pseudorotation* cycle of the five-membered ring. For any given value of P , the sum of all of the positive dihedral angles in the ring must be equal to the sum of all the negative dihedral angles: $z_1 + z_2 + z_3 + z_4 + z_5 = 0$.

One can also write an equation for the out-of-plane coordinate z_j for the atom j , given any value of P and $z_{\max}$:

$$z_j = z_{\max} \sin(144^\circ j + P) \quad (1.1)$$

By definition, $P = 0^\circ$ when the sugar ring of a nucleotide is halfway between the C2'-*exo* (^2E) and the C3'-*endo* (^3E) forms. This is a symmetric twist form, written as ^3_2T , where the 3 and 2 refer to the neighboring *endo* and *exo* conformations (notice that the vertical positions of the numbers is important). Since the pseudorotation cycle resembles a compass, the sugar pucker angles that are close to the ^3_2T form are often called “north” or “*N*.” Similarly, when the sugar ring has a pseudorotation angle of $P = 180^\circ$, it is halfway between the C3'-*exo* (^3E) and the C2'-*endo* (^2E) forms. This is the twist form ^2_3T , which is called “south” or “*S*.”

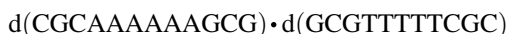
The P values corresponding to different deoxyribose ring conformations can be seen in Fig. 1.12. In this description, the E and T conformations alternate every 18° as one moves around the pseudorotation cycle. Real sugar rings need not fall within these discrete categories, but may exist at any energetically allowed P value. Thus, the conformation of every deoxyribose moiety in a DNA chain can be represented by an angle P . It is noteworthy that these different conformations also have different angles (χ) between the glycosidic bond and the average plane of the deoxyribose ring (see Section 1.1.8).

Different DNA conformations actually have different energetic preferences for their sugar conformations arising from both environmental conditions and DNA sequence. Thus, the conformation C3'-*endo* ($P \sim 0^\circ$, ^3E , or *N*) predominates in A-form DNA, while that of C2'-*endo* ($P \sim 180^\circ$, ^2E , or *S*) is the most common in B-DNA (see Section 1.1.8). Table 1.1 summarizes these comparisons.

1.1.5. Variations of B-DNA Structure

Considerable variations have been found among the structures of B-DNA studied by X-ray diffraction of single crystals of oligodeoxyribonucleotides. A structural study by Dickerson and Drew of the self-complementary dodecamer, d(CGCGAATTCGCG)₂, revealed that this oligomer contains a variety of deoxyribose conformations, although the overall structure is still in the category of B-DNA [17]; it still has a double-helix structure, with a pitch of 35.8 Å, which corresponds to 10.1 base pairs (bp) per turn, and, unlike A-form DNA, all the base-pair planes are nearly perpendicular to the helical axis. However, standard B-form DNA has its sugars in the C2'-*endo* (or “south”) conformation, whereas the deoxyribose conformations found in this structure indicate that they are, in reality, greatly dependent on the nucleotide sequence (see Table 1.3). Higher-resolution studies of this dodecamer confirmed these results [18,19].

An extreme deviation from the B-DNA canonical structure was found by Nelson and others [20], in the “A-tract sequence” of the oligonucleotide duplex:



Here it was found that the extended string of contiguous (non-alternating) A-T base pairs results in an increased “propeller twist” within each base pair. In other words the bases within an A-T base pair are non-planar with respect to each other, and are rotated in opposite directions (contrarotated) along their longitudinal axis, as seen in Fig. 1.13.

TABLE 1.3. Deoxyribose Conformations in the Crystal of DNA Double-Helix Dodecamer, d(CGCGAATTCGCG)₂

Residue	Conformation	Residue	Conformation
C ₁	C2'-endo	C ₁₃	C2'-endo
G ₂	C1'-exo	G ₁₄	C1'-exo
C ₃	O1'-endo	C ₁₅	O1'-endo
G ₄	C2'-endo	G ₁₆	C2'-endo
A ₅	C1'-exo	A ₁₇	C2'-endo
A ₆	C1'-exo	A ₁₈	C2'-endo
T ₇	O1'-endo	T ₁₉	C1'-exo
T ₈	C1'-exo	T ₂₀	C1'-exo
C ₉	C1'-exo	C ₂₁	C1'-exo
G ₁₀	C2'-endo	G ₂₂	C2'-endo
C ₁₁	C2'-endo	C ₂₃	C1'-exo
G ₁₂	C1'-exo	C ₂₄	C3'-endo

Source: Ref. 17.

The average angle of rotation is 20°, which improves the overall stability of the helix by maximizing purine–purine base-stacking interactions. In this case, the helical repeat averages 10.0 bp per turn, and the helical structure is still similar to that of B-DNA. Nevertheless, A-tracts have unusual properties that prevent them from associating into nucleosomes, and short runs cause bending of the DNA.

Many other DNA oligomers have been crystallized and analyzed by X-ray diffraction, some of which are listed in Table 1.4. Those with duplex B-DNA structures are listed in the middle column [21–29]. A more complete review of single crystal X-ray diffraction studies of oligonucleotides is available in Ref. 30.

1.1.6. Participation of Water Molecules in B-DNA Structure

So far, we have considered three factors that stabilize the double-helix B-DNA structure: (1) hydrogen bonding between complementary base pairs, (2) π -electron interaction between the bases stacked vertically, and (3) electrostatic forces between the phosphate group (PO_2^-) and solvated cations (Na^+ , K^+ , Mg^{2+} , etc.). Another indispensable contribution to the stability of the DNA structure is hydration. This is demonstrated by the fact that the sodium salt of a DNA (Na^+ -DNA) fiber produces different X-ray diffraction patterns depending on the relative humidity of its environment. These fibers produce a clear, standard, B-DNA X-ray diffraction pattern only if the fiber is maintained in an environment with 92% relative humidity. When the humidity drops lower than 75%, the X-ray diffraction pattern changes significantly, indicating that the conformation of the DNA fiber has changed to that of A-form DNA. When the humidity drops below 60%, it becomes difficult to obtain a well-defined X-ray diffraction pattern, suggesting that the DNA no longer has a well-defined structure [31].

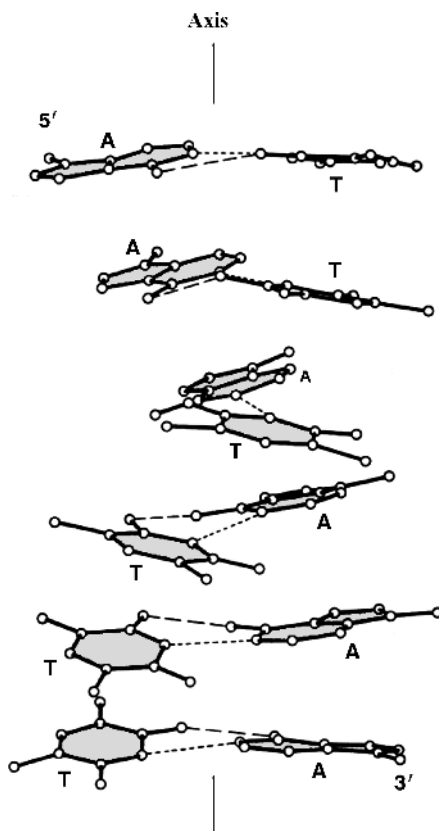


FIGURE 1.13. The central d(A)•d(T) tracts of the structure of double-helix DNA oligomer, d(CGCAAAAAGCG)•d(GCGTTTTTTT CGC) [20]. Only the bases are depicted.

High-resolution X-ray diffraction studies of single crystals of DNA oligomers have also revealed that DNA has a spine of water molecules along its minor groove [18,32]. As shown in Fig. 1.14 for B-DNA, this spine of water molecules has a zigzag shape running along the helix; they alternate between radial positions, where one H₂O molecule is located deep inside the minor groove while the neighboring H₂O molecule is located toward its outer edge. All of these water molecules stabilize the structure through hydrogen bonding to the O and N atoms of the duplex.

In most cases, groove-binding drug molecules such as netropsin [33] replace the spine of water molecules, thereby stabilizing both the duplex structure and the drug–DNA complex.

1.1.7. Partially Dehydrated Structure: A-DNA

In Section 1.1.6 we noted that the sodium salt of a DNA (Na-DNA) fiber adopts a conformation that is largely in the A form when the relative humidity is >75%,

TABLE 1.4. Oligodeoxyribonucleotides Subjected to X-Ray Crystallographic Studies^a

A-DNA	B-DNA	Z-DNA
CCCCGGGG [35]	CGCGAATTCGCG [21]	CGCG [53]
GCCCCGGG [36]	CGCAAAAAGCG [20]	CGCGCG [51]
GGCCGGCC [24]	CGCAAAAATGCG [22]	CGCGCGCG [51]
GGGCGCCC [37]	CGCATATATGCG [23]	CGCATGCG [54]
CTCTAGAG [38]	CGTGAATTCACG [24,24]	CGTACGTACG [55]
GTACGTAC [39]	CGATCGATCG [26]	
ATGCATGC [40]	CCAACGTTGG [27]	
GGTATACC [25]	CCAGGCCTGG [28]	
GGGATCCC [41]	ACCGGCGCCACA [29]	
GGGTACCC [42]		
GGGGCCCC [43]		
GGATGGAG [44]		
GTGTACAC [45]		
ACCGGCCGGT [46]		

^aReference numbers appear in brackets.

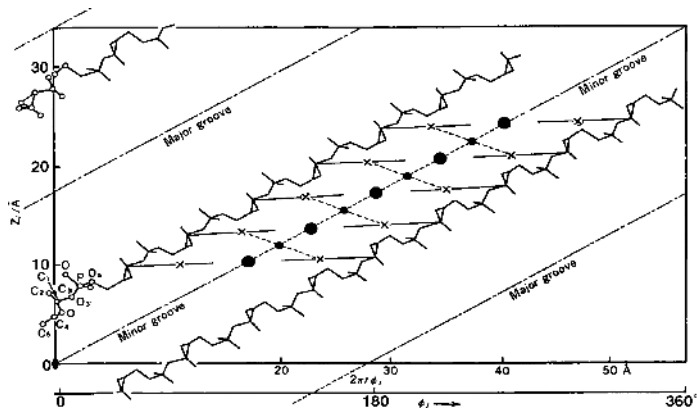


FIGURE 1.14. A spine of water molecules in the minor groove of B-DNA. This is a projection of the B-DNA structure on a plane that is formed by unrolling of a cylinder, that is, constructed by linking all P atoms in the duplex. As given in Table 1.2, every P atom is located at $r = 8.91 \text{ \AA}$ from the helix axis, and therefore the P cylinder should have a circumference of $2\pi r = 56.0 \text{ \AA}$. Thus, the plane for projection should be a rectangle of $33.8 \times 56.0 \text{ \AA}$ as shown here. The projections of polynucleotide mainchains are running from the lower left corner toward the upper right corner. A line of large and small circles is seen, along the minor groove that lies between two parallel polynucleotide chains. The small circles represent water molecules located deep inside the groove; the large circles represent water molecules located closer to the surface of the groove. The deeper water molecules form hydrogen bonds with O and N atoms (shown by “X”) of the base residues (shown by horizontal lines projected out from C1’ atom in the mainchain), whereas the shallow water molecules form hydrogen bonds with the O atoms in the diphosphoester and the deoxyribose ring [18, 52].

and adopts a B-form conformation when the humidity is $\sim 92\%$. That the A and B forms are different conformations of the same DNA molecule is confirmed by the observation that changes in humidity induce reversible changes between A and B conformations [31].

As does B-form DNA, the A-form structure [34] (Fig. 1.9a) has a double-stranded structure consisting of two interwound right-handed helices, which run antiparallel to each other. In contrast to B-DNA, however, the midpoint between the base-pairs is located outside the center of the helical axis, and have an inclination angle of $\sim 19^\circ$. The overall structure of A-DNA is shorter, fatter, and more compact than that of B-DNA. The helical pitch of A-DNA ($\sim 28 \text{ \AA}$) is significantly less than that of B-DNA ($\sim 36 \text{ \AA}$), having 11 nucleotide residues per turn instead of 10. The rise per base pair is much shorter ($\sim 2.6 \text{ \AA}$) compared to B-form DNA ($\sim 3.4 \text{ \AA}$). An important indicator of A-form DNA is that its deoxyribose conformation is in the $C3'$ -*endo* (*N*) conformation, instead of the $C2'$ -*endo* (*S*) conformation found in canonical B-DNA. These differences are summarized in Table 1.1.

Many relatively short DNA oligomers have crystallized into the A form, as determined by X-ray diffraction studies. Some of these A-DNA structures are listed in the first column of Table 1.4 [24,25,35–46].

In solution, a long stretch of A-DNA can be obtained by adding ethanol (85% in volume) to the aqueous solution of long B-DNA. The biological significance of A-DNA has not yet been clearly established, as so far there is no direct evidence for its presence in a living system. It should be pointed out, however, that double-helix RNA is structurally very similar to A-DNA, and that DNA–RNA hybrids have A-DNA-like structures even in aqueous solution without ethanol. In addition, the A form may be induced when DNA binds with certain drugs or proteins.

1.1.8. Internal Dihedral Angles

The bond lengths and bond angles in DNA remain nearly unchanged even when there is a significant conformational change caused by the binding of a drug or ligand. Instead, the conformational changes are attributable to changes in some of the dihedral angles within the DNA structure. As shown in the upper left corner of Fig. 1.15, every deoxyribonucleotide unit has seven single bonds, around which the atomic groups at both ends can rotate nearly freely (across lower potential energy barriers). These dihedral angles are defined as follows. If a system of four atoms A–B–C–D is projected onto a plane perpendicular to (“normal to”) the B–C, then the angle between the projection of A–B and that of C–D is described as the dihedral angle about the B–C bond. The value of the dihedral angle is considered to be positive, if a clockwise turn of the A–B bond about the central bond is necessary to eclipse it on the C–D bond. The value is negative if a counterclockwise turn of the A–B bond is needed to eclipse it on the C–D bond (upper right corner of Fig. 1.15). It is convenient to remember that this definition can be rephrased as follows. If a clockwise turn of the bond containing the front atom about the central bond is needed for it to eclipse the bond to the rear

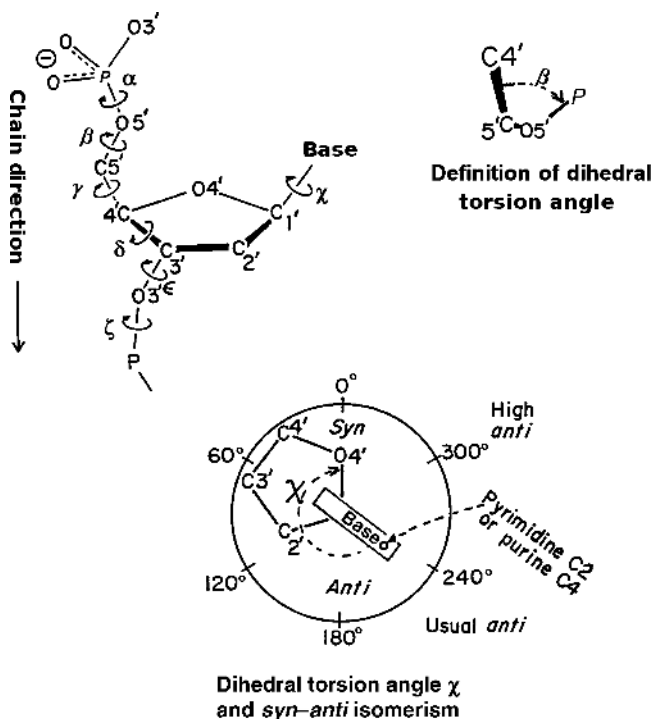


FIGURE 1.15. Internal torsion angles for expressing a local conformation of deoxyribonucleotide chain.

atom, then the angle is positive, regardless of the end from which the system is viewed [47].

A histogram of the internal dihedral angles within a DNA chain can be constructed [48] with reference to the crystallographic data from several oligodeoxyribonucleotides [24,25,49], and the results are shown in Fig. 1.16. The dihedral angles α , β , and γ are approximately -60° , 180° , and 60° , respectively, for both B- and A-DNA conformations. However, B-DNA and A-DNA are distinctly different conformers, because δ , ϵ , and ζ are different. Thus, $\delta = 120^\circ \pm 40$ for B, but $\delta = 90^\circ$ for A; $\epsilon = -180^\circ$ for B, but -150° for A; and $\zeta = -90^\circ$ for B and -80° for A. The torsion angle δ is related to the pseudorotation phase-angle P of the five-membered ring $C1'-C2'-C3'-C4'-O4'$, as shown in Fig. 1.17. Thus, the distribution of δ values in B-DNA (shown in Fig. 1.16, lower) indicates that its deoxyribose ring can adopt a wide range of P values, as pointed out by Dickerson et al. [49]. The ζ distribution has a tail toward -60° in A, and toward -180° in B. This unusual distribution of the dihedral angle ζ seems to be related to a sequence-dependent conformational variation. Similar, but much more detailed, histograms of the torsion angles have been obtained by high-resolution structural studies of B-DNA, A-DNA, and Z-DNA [50].

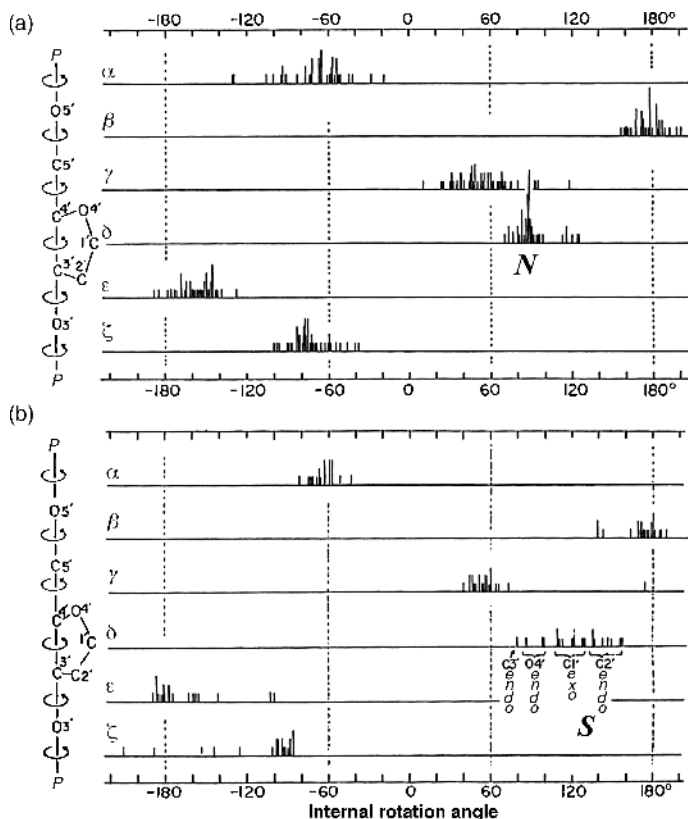


FIGURE 1.16. Histograms of the torsion angles in single crystals of double-helix DNA fragments: (a) d(GGCCGGCC)₂ [24] and d(GGTATACC)₂ [25]; (b) d(CGCGAATTCGCG)₂ [21].

The dihedral angle about the *N*-glycosidic bond (C1'–N), which links the base to the sugar, is denoted by the symbol χ . The definition of χ is illustrated in the lower right corner of Fig. 1.15. The sequence of atoms chosen to define this angle is O4'–C1'–N9–C4 for purine and O4'–C1'–N1–C2 for pyrimidine derivatives. Thus, when $\chi = 0^\circ$, the O4'–C1' bond is eclipsed with the N9–C4 bond in purines and eclipsed with the N1–C2 bond in pyrimidines. This conformation is called *syn*. A conformation with $\chi \gg 0^\circ$ is called *anti*. In general, the glycosidic bonds in B-DNA have an *anti* conformation.

1.1.9. Left-Handed Duplex: Z-DNA

Yet another type of duplex DNA conformation is that of Z-DNA. This duplex DNA structure is quite different from the preceding forms, A and B, in that it is a left-handed helix, rather than a right-handed one. This means that if one imagines

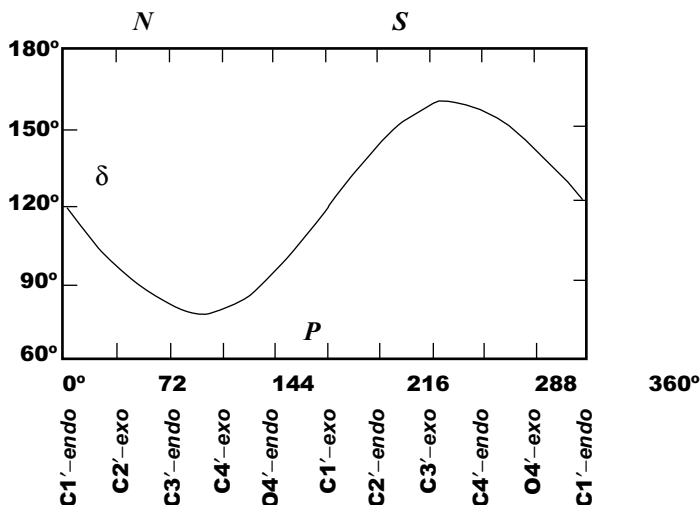


FIGURE 1.17. Relation between the C5'–C4'–C3'–O3' internalrotation angle δ and the pseudorotation phase angle P of the deoxyribose ring.

grabbing a Z-DNA strand with one's left hand, with the thumb pointing in the direction of the helical axis, then the twist of the helix will be in the same direction as the curve of the fingers. One of the earliest reports on this conformation is the crystal structure study of the self-complementary oligo-DNA chain, d(CGCGCG)₂ [51,52]. This duplex structure (Fig. 1.9c) has the standard Watson–Crick, G–C base pairing, but the glycosidic bond angles (χ) of the guanine bases adopt the *syn* conformation, while the cytosine remain in the standard *anti* conformation. Since the sequence alternates between G and C, their attached sugars also alternate between *syn* and *anti*, as one moves along the strand. The deoxyguanosine sugar residues are also unusual in that they are in the C3'-*endo* (*N*) conformation, while the deoxycytidine residues are in the C2'-*endo* (*S*) conformation. (See Table 1.1 for a summary of these comparisons.) Some of other DNA oligomer sequences that have been found to have the Z-DNA structure by X-ray crystallographic investigations [53–55] are listed in the third column of Table 1.4.

Numerous studies have shown that poly(dG–dC)₂ can adopt a Z-DNA conformation in solution and that this conformation is stabilized by several chemical modifications of the base residues [56]. The Z-DNA conformation can also be adopted by the family of DNAs, poly[d(A–C)] • poly[d(G–U)], which bear at least one methyl or halogen (Br or I) at the pyrimidine C5 heterocyclic position [57]. Also, many G4-quadruplex structures (RNA and DNA) adopt Z-like conformations in which the guanines have alternating *syn/anti* glycosidic bond angles (χ) along the strands (Section 1.1.12).

Ha and others [58] constructed a synthetic DNA structure that contained a junction between B-DNA and Z-DNA. A 15-bp DNA duplex was designed such that it had an additional two nucleotides at the 5' end of each strand that were not base-paired, and which overhung the duplex formed by 5'-d(GTC₋₈G₋₇C₋₆G₋₅C₋₄G₋₃C₋₂C₋₁A₀T₁A₂A₃A₄C₅C₆) - 3' and 5'-d(ACG₋₆G₋₅T₋₄T₋₃T₋₂A₋₁T₀G₁G₂C₃G₄C₅G₆C₇G₈) - 3'.

The alternating dG-dC sequence composed the Z-DNA portion of the structure, and the other half of the duplex formed B-DNA. In this case, the bases were numbered from the center junction, with the two strands differentiated by the prime (') notation. The Z-DNA structure did not stabilize on its own; instead, these workers took advantage of the tight binding of a Z-DNA binding protein to induce this conformation in one half of the DNA duplex, while the remainder of the DNA duplex retained its B conformation. One B-Z junction was thus formed in the middle of the DNA duplex, connecting the Z- and B-DNA. Their X-ray crystallographic study indicated that the structure of the 8-bp region in the Z conformation is essentially the same as the Z-DNA structure described earlier (Fig. 1.9c), with a zigzag sugar phosphate backbone and alternating *anti-syn* base conformation. The other end of the DNA duplex adopts a standard B conformation (Fig. 1.9b). Between the B and Z conformations, an A-T base pair is disrupted and one base from the center of each strand (T₀' and A₀) is extended out from the double helix (Fig. 1.18).

The glycosyl bonds of A₀ and T₀' are both in the *anti* conformation. The base of T₀' is oriented almost parallel to the helix axis, whereas the A₀ base is fully extended out from the helix (Fig. 1.18). Thus, the handedness of the DNA duplex can be completely reversed merely by breaking one base pair and extruding the bases from the duplex. However, Z-DNA is inherently unstable, and to form Z-DNA under physiological conditions requires additional energy, such as can be supplied by negative supercoiling of the DNA (see Section 1.1.10). The base extrusion at the B-Z junction is accompanied by a sharp turn in the sugar-phosphate backbone, a slight bend (11°) at the junction, and displacement of the two helical axes by 5.2 Å. Thus, with only a single base pair broken at the junction, the structure affords a remarkable conservation of energy and maintenance of helical structure.

1.1.10. Superhelical Form of DNA

The tertiary structure of DNA includes superhelices of the double-stranded DNA chain. A *superhelix* is a helix formed by the curvature of the axis of the DNA duplex. The axis of B-DNA, for example, appears as a straight line in terms of its local structure, but in the cell, it is actually slightly curved, such that in a long strand of B-DNA duplex, it is no longer straight but forms another helix. Hence, in essence, a superhelix is a helix of a helix. A superhelix is often produced by a slight unwinding of the duplex, which may be induced by drug binding. In the canonical B-DNA, a base pair can be superimposed upon an adjacent one by translating it 35.5 Å along the helical axis, while simultaneously rotating it 36° around the helical axis (since there are 360° in

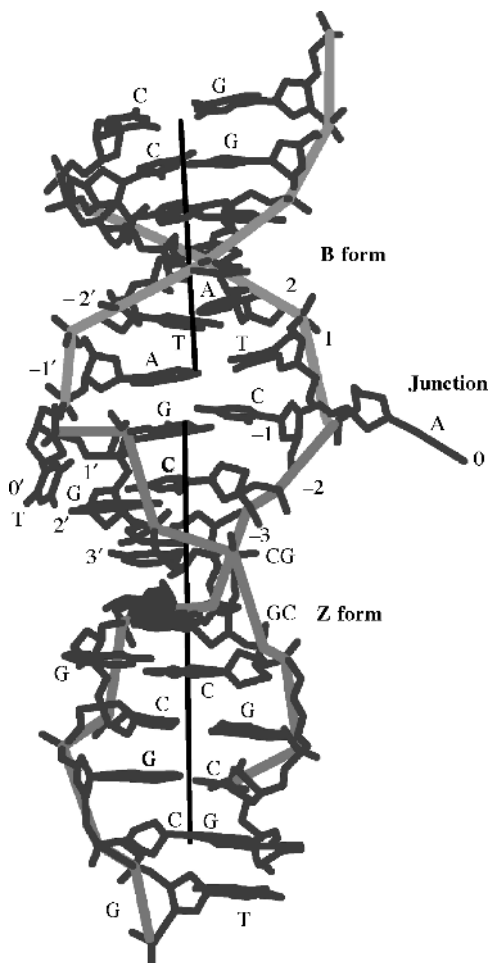


FIGURE 1.18. A 15-base-pair (15-bp) DNA structure containing a junction (black) between left-handed Z-DNA (magenta) and right-handed B-DNA (cyan) [58]. (See color insert.)

a complete turn and ~ 10 bases per turn; hence $360^\circ/10$). If, for example, the structure is deformed so that this angle is reduced from 36° to 31° , then the unwinding angle is 5° .

Superhelix formation often takes place in closed, circular DNA duplexes, as in *E. coli*, where the 3' end and 5' head of each strand are covalently connected through a phosphodiester linkage (Fig. 1.19). If a relaxed string of B-form DNA is closed into a circle, there will be no supercoiling (Fig. 1.19a). However, if the DNA double helix is unwound by 360° prior to joining its ends (Fig. 1.19b), it is harder for it to adopt the B form, even though it is energetically favorable to do so. The B form can be attained, however, by “winding up,” or tightening, by 360° . Since the DNA is enclosed within a circle, this tightening introduces a strain, which is compensated by winding the circle of DNA into a superhelix with one turn (see Fig. 1.19c).

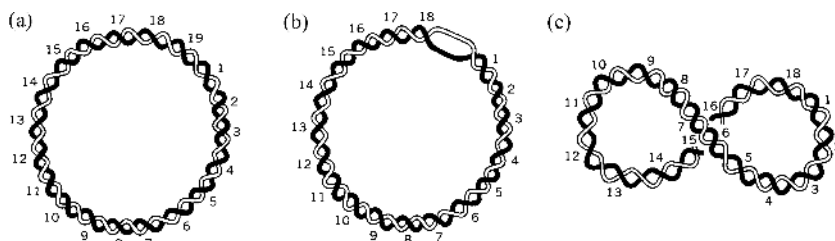


FIGURE 1.19. Schematic illustrations of closed circular DNA duplex. (a) Relaxed form, where linking number (α) = 19, twisting number (β) = 19, and writhing number (τ) = 0 (no supercoiling). The figures along the chain indicate the number of times one strand (black chain) goes around the other (white chain). (b) A hypothetical form that might be produced immediately after the circle is closed with the linking number (α) = 18, instead of 19. The closing portion does not have the B-DNA conformation and must be very unstable. (c) The unstable portion has been brought into a stable B-DNA structure, by twisting the whole circle by one turn, producing a negative superhelix.

In general, the topological properties of closed circular duplex DNA are defined by the relationship [59,60],

$$\alpha = \beta + \tau \quad (1.2)$$

where α is the “linking number” of the closed circular duplex, which is the number of times one strand goes around the other, and is a fixed integral number; β is the “twisting number,” which is the number of helical turns of the DNA duplex; and τ is the “writhing number,” which is the number of superhelical turns. In the model shown in Fig. 1.19a, $\alpha = 19$, $\beta = 19$, and $\tau = 0$. The conformation given in Fig. 1.19b does not actually exist. For the conformation shown in Fig. 1.19c, $\alpha = 18$, $\beta = 19$, and $\tau = -1$. The meaning of $\tau = -1$ is that the DNA has been unwound by one turn.

Natural, closed, circular DNAs are much larger than the model shown in Fig. 1.19. For example, consider the intact viral plasmid pBR322, which is purified from *Escherichia coli*, and is a closed circular DNA duplex consisting of 4361 bp. It is known to have a highly negative supercoiling ($\tau < -20$). By the action of topoisomerase I, however, it can be brought into the relaxed B form ($\alpha = 428$, $\beta = 428$, and $\tau = 0$). This was demonstrated by the preparation of moderately supercoiled pBR322 closed circles (e.g., $\alpha = 420$, $\beta = 428$, and $\tau = -8$) using the addition of appropriate amounts of aclacinomycin, followed by the addition of topoisomerase [61].

1.1.11. Other Duplex Structures

1.1.11.1. Mismatched Base Pairs Errors during replication or repair of DNA can lead to the presence of mismatched base pairs or unpaired bases within the double helix. For example, a mismatched base pair was found by X-ray crystallography [62] in the dodecamer duplex d(CGCAAATTCGCG)₂. Notice that this 12-base sequence is almost self-complementary—the string of Ts is one base shorter than the string of As—and it will prefer to exist in solution as a hydrogen-bonded duplex, even though the A-C base pairing at the ninth position is not correct. Figure 1.20 shows the A-C base

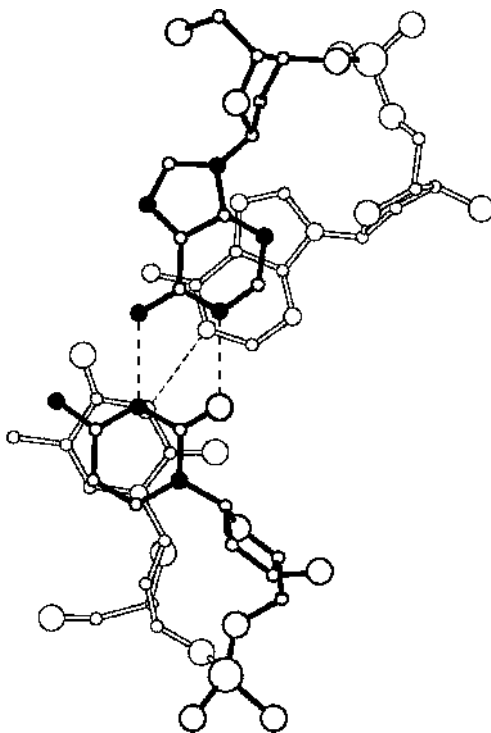
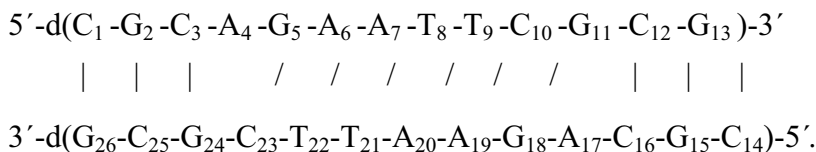


FIGURE 1.20. The A-C base pair, seen in bold lines, as established by the X-ray analysis [62]. This is proposed as a model of the A-C mismatch in double-helical DNA.

pair observed in the crystal structure of this dodecamer. The probability of incorporating an incorrect base during replication and its subsequent detection and excision by polymerase is thought to be related to the thermodynamic stability of the particular base-pair mismatch in the double helix. The relative energetic stabilities of different mismatches were studied by NMR (nuclear magnetic resonance) in a series of sequence-related dodecamers by determining their midpoints in the helix-coil transition. Relative to the fully hydrogen-bonded duplex form of the initial dodecamer, it was found that the G-T and A-G base pairs destabilize the double helix by 17–20°C, while the A-C base pair destabilizes it by 30°C [63]. Thus, even a single mismatch results in a significant lowering of the overall stability of a duplex, making it a looser structure. The stability of this short duplex is increased by the terminal G-C base pairs, each of which has three hydrogen bonds, and as such are said to be “stickier.” As the length of the duplex portion of the DNA increases, this destabilization will contribute less on a percentage basis to the total energy. This becomes important in biological systems, where the native DNA is many thousands of bases long, and as such is capable of withstanding several mismatches.

1.1.11.2. Bulged Double Helix Sometimes two strands of DNA are complementary, with the exception of a stretch of extra bases in the middle of one strand.

These extra bases cannot form base pairs, and hence are bulged out of the duplex. The 13-base oligomer d(CGCAGAATTCGCG) is an example a nearly self-complementary sequence:



When the two strands of this sequence come together to form a duplex, the A at the fourth position might be expected to form a mismatched base pair with the C on the other strand (e.g., A₄-C₂₃). This is not a Watson-Crick base pair, however, and instead the two additional adenosines (A₄, and its equivalent, A₁₇) are looped out from the double helix, forming a small bulge (Fig. 1.21) [64]. The looped-out bases cause little

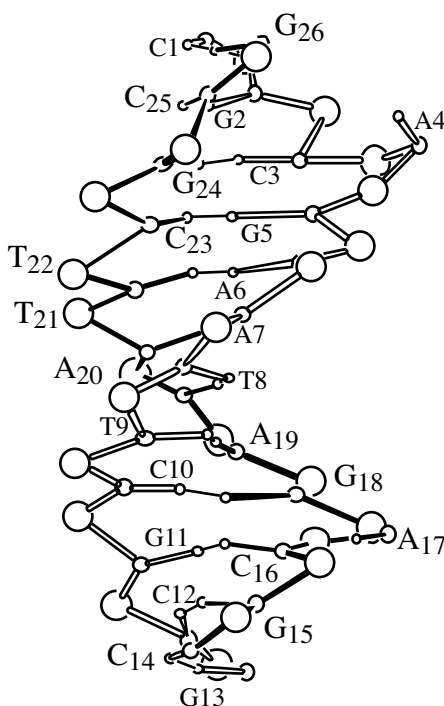


FIGURE 1.21. A schematic representation of the structure of the tridecamer d(CGCA-GAATTCGCG)₂ structure showing the extra adenosines (A₄ and A₁₇) bulging out [64]. The sugar-phosphate backbone is represented by phosphorus atoms (large spheres) linked to C1' atoms of sugars (medium-sized spheres), and the bases by links from the C1' atoms to the N9 atoms in purines or to the N1 atoms in pyrimidines (small spheres). The base-pair interactions are drawn as thin lines.

disruption of the rest of the structure, and the overall conformation is very close to that of B-form DNA.

1.1.11.3. Hairpin Double Helix A DNA duplex does not need to be composed of two separate strands. When the DNA strand concentration is low, an otherwise self-complementary sequence can fold upon itself and adopt a monomeric hairpin conformation. This can also happen with a sequence that is only partially self-complementary. One example is the synthetic, 16-base DNA oligomer d(CGCGCGTTTTTCGCGCG). A crystal structure of this sequence was solved at 2.1 Å resolution and was observed to form a monomeric hairpin having a Z-DNA hexamer stem and a loop consisting of four thymine bases (called a “T4” loop) (Fig. 1.22) [65]. The bases in the loop stack with one another, rather than with the base pairs of the hexamer stem. A similar stacked-base arrangement was proposed for the T4 loop of d(CGCGCGTTTTTCGCGCG), on the basis of NMR measurements in solution and distance geometry calculations [66] (Section 1.7.2.1). In that case, however, the hairpin has a B-DNA stem, and the thymine bases in the loop were stacked parallel to the final base pair on the top of the stem. The difference in the T4 loop conformation between the X-ray and NMR studies probably reflects differences between crystal and solution environments, and illustrates the flexibility of a four-base loop. In general, it has been found that the details of the

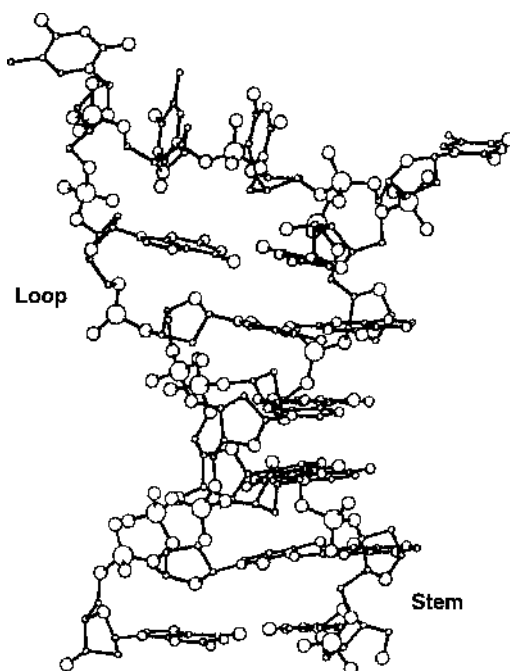


FIGURE 1.22. A structure of the hairpin of d(CGCGCGTTTTTCGCGCG) [65]. Atoms in decreasing size are P, O, N, C. The stem of the hairpin is a normal left-handed Z-DNA hexamer.

stacking of bases within a loop is correlated to the adjacent base pairs at the top of the stem, with some stabilizing stacking more than others.

Two nearly self-complementary strands can also come together to form a duplex with an internal loop. In this case, there is a loop, or bulge, in the middle of both strands.

1.1.12. Higher-Order DNA Structures

The polymorphism of DNA has become a subject of considerable study in recent years, particularly because it has been suggested as a means of gene regulation and control of gene-based disorders and diseases. These non-B-DNA conformations include intra- and intermolecular triplex DNA, quadruplex DNA, cruciforms, and slipped mispaired structures. The first two of these are briefly discussed below, owing to their more unusual hydrogen-bonding patterns. For a variety of other DNA structures, see Ref. [67], for example.

1.1.12.1. Triple Helix Triple-helix formation between a homopyrimidine–homopurine duplex and a third homopyrimidine strand was discovered by Felsenfeld and Rich [68] in 1957, but remained a curiosity for many years. A possible biological role for triplex DNA has emerged in the area of gene regeneration with the discovery of H-DNA, which is a DNA structure formed by a pair of mirror repeat duplex sequences that contains a triple helix region, as seen in Fig. 1.23 (top) [69,70]. Triplex formation (including H-DNA formation) has received attention as a possible means to artificially control gene expression [71]. The idea here is that a third strand could be designed to bind in a sequence-specific manner to the duplex DNA of a particular gene and inhibit it; hence it is sometimes called an “antigene” approach. This approach, which involves the use of a piece of DNA as if it were a drug, could be ideal in terms of sequence specificity. It has faced two important limitations: (1) only some disease states have DNA target sequences that can form a triplex; and (2) it is difficult to deliver these short oligonucleotides to the correct gene in the cell without being digested by nucleases.

Triple-stranded DNA may be formed by laying a third strand into the major groove of a DNA duplex. Complementary hydrogen bonding interactions are responsible for the specificity of the third strand interaction. In these structures, the Watson–Crick base pairing within the duplex is unchanged, while the third strand is hydrogen-bonded to the duplex via a Hoogsteen base-pairing scheme. The hydrogen bonding patterns within in triple-stranded DNA are called “base-triples,” and their schemes are *pur–pur–pyr* (GGC and AAT) and *pyr–pur–pyr* (TAT, CGC), the latter of which is shown in Fig. 1.23 (bottom). The central strand of the triplex must be purine-rich since pyrimidines do not have any additional hydrogen-bonding surfaces. Thus, triple-stranded DNA requires a homopurine–homopyrimidine sequence that is at least several bases long. If the third strand is purine-rich, it forms reverse Hoogsteen hydrogen bonds in an antiparallel orientation with the purine strand of the Watson–Crick duplex. If the third strand is pyrimidine-rich, it forms Hoogsteen bonds in a parallel orientation with the Watson–Crick-paired purine strand (Fig. 1.23, bottom).

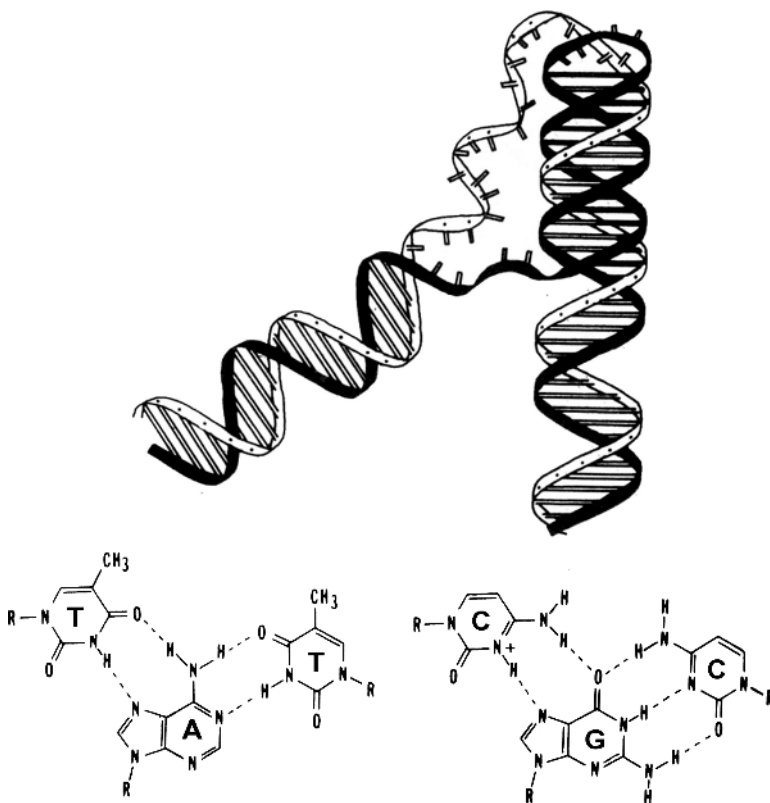


FIGURE 1.23. Structure of intramolecular triplex [70]. The strand containing the pyrimidine-rich sequence is shaded, whereas the strand containing the purine-rich sequences is unshaded. The 3' half of the purine strand of the insert remains Watson–Crick-paired with the 5' half of the pyrimidine strand. The 3' half of the pyrimidine strand, after dissociation from its Watson–Crick complement, occupies the major groove of the duplex region described above and is specifically complexed with it through Hoogsteen base pairs. The mirror repeat and the oligopurine–oligopyrimidine nature of these sequences enable this specific association.

1.1.12.2. Quadruplex DNA Four strands of DNA may interact to form a stable structure in a number of ways, and evidence for the importance of these structures *in vivo* is increasing. Similar sequences have been found, for example, in the telomeric DNA at the ends of chromosomes, suggesting that quadruplexes play a role in cell replication, and hence are likely to be important in both cancer and the aging process [72–76]. Involvement of quadruplexes in the promoter regions of oncogenes [77] and in the dimerization of the HIV-1 retrovirus [78] has been proposed. They also have been implicated in the regulation of the insulin gene [79], in the triplet repeat sequences associated with certain neurological disorders, such as the “Fragile X syndrome” [80–83], and in autoimmune diseases such as erthemaosus lupus [84]. Similarly, there is also a growing list of proteins that appear to bind specifically

to quadruplex structures [85–90]. Some proteins seem to facilitate their formation, whereas the activity of other proteins appears to be suppressed by these structures.

Because of the prevalence of those quadruplex structures *in vivo*, they have been proposed as drug targets, and this has stimulated interest in the possibility of a whole new category of drug function. The primary action of most drugs is to bind to DNA, thereby interfering with transcription or replication. However, these proposed drugs would operate by altering the tertiary structure of the DNA or RNA, inducing gross structural changes through the enhanced formation of quadruplexes, or the destabilization of existing ones. For example, controlling the formation of G-quadruplex structures (discussed below) in telomeric DNA has been proposed as a means of treating cancer [91], while destabilizing them in triplet repeat sequences might aid in the treatment of other diseases. The binding mechanisms of these drugs normally follow the same general principles as has been determined for duplex DNA, and are not treated separately in this book.

The unique structures and properties of quadruplex DNA require some additional explanation. Instead of “base pairs,” quadruplexes have “base-quartets” consisting of four bases that are hydrogen-bonded to each other. The most common form of quadruplex is based on the guanine quartet, and has been recognized since 1962. Optical spectroscopic studies using infrared (IR) and circular dichroism (CD) demonstrated that individual guanines, and its derivatives, can self-associate in the presence of cations to form Hoogsteen-type, hydrogen-bonded G-quartets that can then be stacked vertically (on top of each other) [92–95]. Under physiological conditions, guanine-rich sequences of DNA or RNA readily form these structures *in vitro*. Such guanine-rich structures have been variously named “G-quadruplex,” “G4-quadruplex” or “G4-DNA.” Quadruplexes may be composed of four separate strands that have associated with each other. In this case, the structure is called “linear,” and all of the DNA strands are aligned in a parallel manner (Fig. 1.24a). Quadruplexes can also be formed when one or more DNA strands fold upon themselves. They can have as many as three folds per structure. Each fold results in a loop region (e.g., see Fig. 1.24b), and usually results in the reversal of the strand polarity in the core of the quadruplex. In other words, within the guanosine quartet region of a G-quadruplex, the polarity of one end of the folded strand is usually antiparallel to its neighboring strands. The arrangement of the strand polarities within these folded G-quadruplexes is sensitive to the sequence, as well as other experimental conditions.

Quadruplexes have considerable variation in their structures, depending, in part, on the length and constituents of any adjacent loop regions. When a G-quadruplex is formed with all of its strands aligned in the linear, all-parallel manner, then all of the guanosine glycosidic dihedral angles are typically in the *anti* conformation (Fig. 1.24a), just as they are in A- and B-form DNA. One notable exception to this was found in the early work by Henderson et al. [96]. However, when the strands are aligned in an antiparallel manner (as in a folded structure), a stable quartet can be formed only if a proportional number of guanosines have glycosidic bond angles that are in the *syn* conformation (e.g., see Fig. 1.24b). In most folded (and therefore, antiparallel) G-quadruplexes,

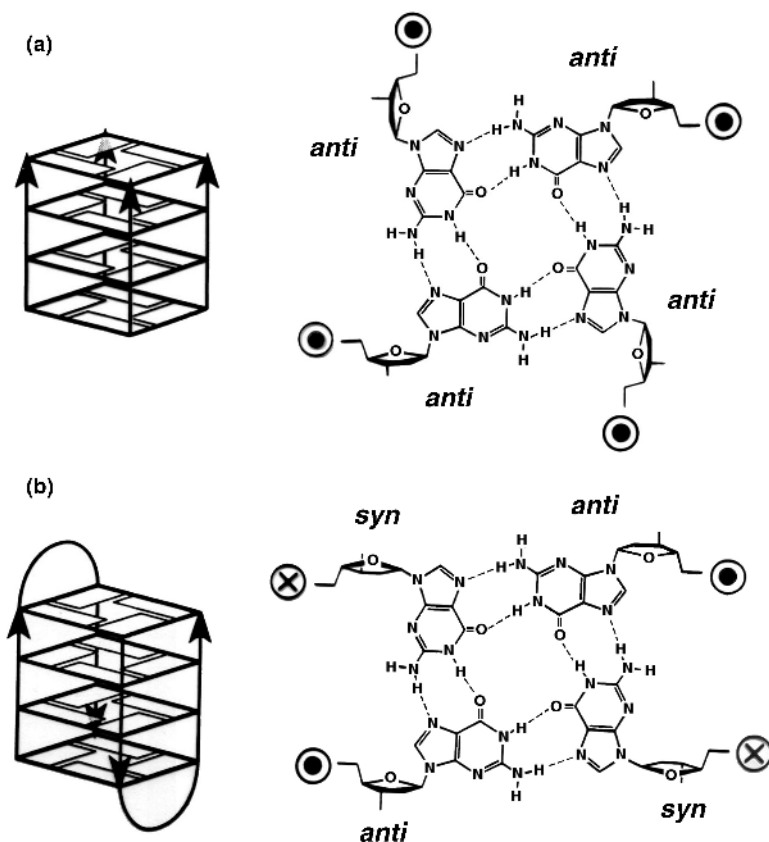


FIGURE 1.24. The core of the quadruplex consists of four guanosine bases arranged in a square, each participating in Hoogsteen base pairing with both of its neighbors. There are a total of eight hydrogen bonds in each "quartet" or "tetrad." The orientations of the strands is illustrated in the side-view (left side of figure) and by the symbols (circle and X) on the right side of the figure. (a) A linear quadruplex in which all strands are parallel and all *N*-glycosidic torsion angles are in the *anti* conformation. (b) A quadruplex arising from strands that have folded into the "edge"-type arrangement. In this arrangement, all of the adjacent strands are antiparallel to each other, and the glycosidic torsion angles alternate between *anti* and *syn*, as one proceeds clockwise around the quartet. These torsion angles also alternate as one moves along the DNA strand.

these glycosidic angles alternate between *anti* and *syn* along the strand (although this is not always the case). This alternation of glycosidic dihedral angles is similar to that of Z-DNA.

Within the G-quartets, many different arrangements are possible for *syn* and *anti* dihedral angles. They may be all *syn* [95], all *anti* [97–100], alternating *syn* and *anti* (Fig. 1.24) [101,102], or *syn-syn-anti-anti* [103–105]. The specific arrangements of *syn* and *anti* dihedral angles also affect the groove widths. Hence, although a base

quartet may appear to have fourfold symmetry, the overall molecule may not have any symmetry at all.

As with duplex and hairpin loop DNA structures, base-stacking interactions are important to the stability of quadruplexes. This includes base stacking within the guanine core [106,107], or between bases in the core and those in a loop or overhanging region [107–109]. Base-stacking interactions within the guanine core are also affected by the glycosidic dihedral bond angles, as well as by the orientations of strands [104].

The DNA sequence can affect the folding pattern of a G-quadruplex structure. Sequences that prefer linear topologies are usually short oligomers containing a stretch of several guanines, such as $d(G_n)$ or $d(G_nT_m)$, where n and m are integers (n is typically 2–5, and m can be almost any number). Longer DNA sequences containing repeating strings of guanines, interspersed with loop segments (consisting mostly of thymines), generally prefer to form folded structures, but they can also form long linear structures. These latter sequences usually take the form of $d(G_nX_mG_n)$, or $d[(G_nX_m)_4]$, where X represents the bases in the loops (mostly T), and both n and m are typically 2–5. These sequences are similar to those found in telomeric DNA. The concentration of the DNA also affects the topology through the principle of mass action. High DNA concentrations favor the four-stranded (linear) species, while lower concentrations favor uni- or bimolecular (folded) species. An equilibrium exists between these forms, and they are often found together in solution.

Critical to the stability of a G-quadruplex structure is the presence of internal, coordinated, cations. The most common cations are the physiologically relevant K^+ and Na^+ . Generally, they are sandwiched in the cavity between adjacent quartets (Fig. 1.25). These cations stabilize the G-quadruplex by electrostatically coordinating with the O6 atoms of the eight guanine bases in the quartet stack, and by contributing favorable free energies of dehydration [110–112]. The K^+ ion has a snug fit inside this

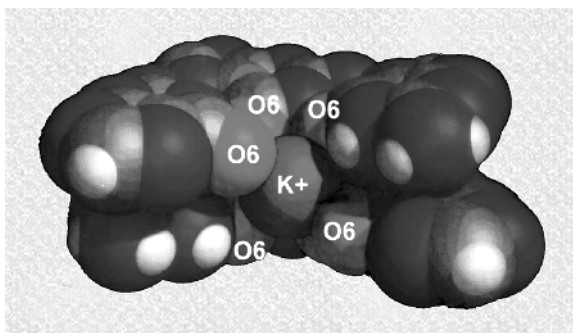


FIGURE 1.25. A side view of two G quartets stacked on top of each other, with an internal, coordinated cation in the center. The view of the stack is from one corner of the square, with the two closest guanosine bases removed to allow visualization of the core. The atom in the center is a K^+ ion, which can be seen coordinating with the O6 atoms in each of the surrounding bases. This figure is derived from experimental NMR structures determined by a combined approach using molecular modeling and experimental NMR data [108]. (See color insert.)

central core, where it touches all eight O6 atoms. The Na^+ ion is smaller and has been sometimes found within the quartet plane. In this case, the Na^+ ion coordinates with only four guanosine bases. The structures stabilized by K^+ are generally more stable than those stabilized by Na^+ , as indicated by the fact that they have higher T_m values (Section 1.2.1). This is consistent with both the more optimal fit and the more favorable dehydration energy of K^+ [111]. Thus, the role of cations in quadruplex DNA is quite different from that in duplex DNA, where they neutralize the charges on the phosphodiester backbone along the outside of the double helix. However, several workers have also suggested that some G-quadruplex topologies may derive additional stability from either groove- or loop-bound cations [113,114].

The cations also preferentially stabilize certain G-quadruplex topologies [115–125]. It has been found that the presence of K^+ tends to stabilize linear quadruplex structures, whereas the presence of Na^+ tends to stabilize structures with folded topologies. This is likely related to the size and location of the cations, but is still a subject of controversy. An interesting observation was that the equilibrium and interconversion pathway between folded and linear topologies may be kinetically blocked for some Na^+ -stabilized structures, preventing it from achieving a thermodynamically preferred configuration [108]. This same equilibrium appears to occur more easily for similar K^+ -stabilized structures. The evidence suggests that the K^+ ion allows for a greater range of molecular motion within the quadruplex, thereby opening up more dissociation pathways for its folded structures, relative to those of Na^+ . Such a kinetically controlled process could be associated with differing degrees of freedom, with the extent of molecular motion depending on the size of the internal cation. This also suggests that it may be possible to develop drugs that can control the specific quadruplex topology that is formed, or even act as another kind of conformational switch.

1.2. ELECTRONIC SPECTRA

1.2.1. Ultraviolet Absorption Spectra

The absorption spectra of five mononucleotides at two different pH values are shown in Fig. 1.26 [62]. These nucleotides have strong bands near 260 nm. At the maxima of their peaks (λ_{max}), the molar extinction coefficients (ϵ) are as great as 10,000–15,000 $\text{M}^{-1} \text{cm}^{-1}$ and their transition moments are in the base planes. These are attributed to $\pi \rightarrow \pi^*$ transitions of the base residues. In addition, another $\pi \rightarrow \pi^*$ transition is observed around 210 nm.

A typical absorption spectrum of a B-DNA solution at room temperature is shown in Fig. 1.27a [126]. It has an absorption maximum at 260 nm and a trough at 230 nm. The 260 nm peak is considered to be caused by $\pi \rightarrow \pi^*$ transitions of the base residues. Such a spectral feature, however, depends only slightly on the base composition and the base sequence of DNA.

The molar extinction coefficient for a DNA nucleotide, at the 260 nm absorption maximum, is always about 6000 $\text{M}^{-1} \text{cm}^{-1}$ (where M = mole/liter per nucleotide), as

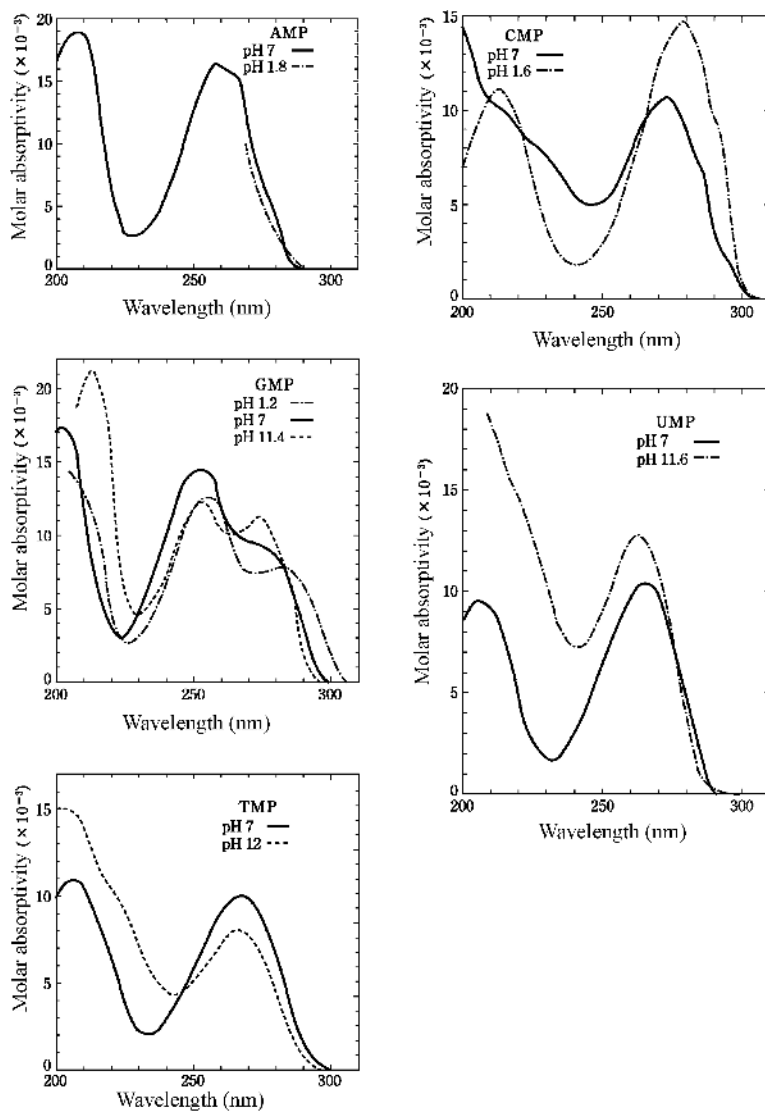


FIGURE 1.26. Absorption spectra of the nucleotides, at 77 K in ethylene glycol–water glass [126].

long as it is in B-DNA conformation [127]. When the duplex is separated into single strands of DNA, such as by heating or some other means, the absorption intensity at 260 nm is greatly increased (see Fig. 1.27a), as the molar extinction coefficient at 260 nm of any mononucleotide far exceeds that of a nucleotide in B-DNA duplex. The peculiarly smaller extinction coefficient of B-DNA results from the interaction of the π electrons between the base aromatic residues when they are stacked vertically.

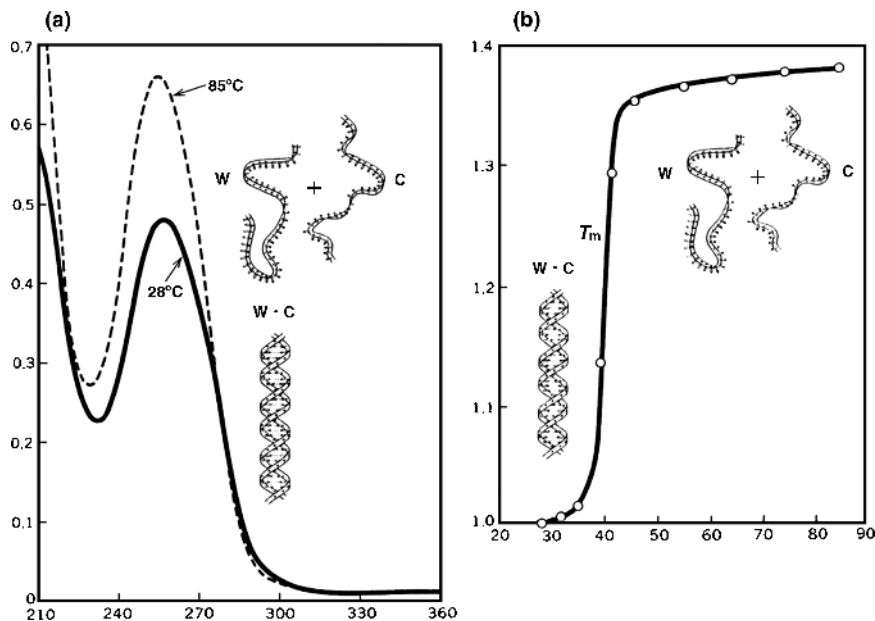


FIGURE 1.27. (a) Absorption spectra of DNA from a crab (*Cancer borealis*) in aqueous solution observed at two different temperatures. Concentration, 7×10^{-5} M; solvent, 0.005 M phosphate buffer, pH = 7. The guanine • cytosine content of this DNA is only 3% [128]. At 28°C, two polydeoxyribonucleotide chains (Watson–Crick chains) are interwound with each other to form a duplex, whereas at 85°C they are separated from each other. (b) A melting curve of the same DNA. In the course of an elevation of the temperature of the solution from 28°C to 82°C, a sharp rise in absorbance at 260 nm takes place at 40°C ($=T_m$, melting temperature).

This decrease in absorption intensity is called *hypochromism*, and its theoretical treatment was described by Tinoco [128].

The intensity of the absorption band at 260 nm can be used to measure DNA concentrations in solution as low as a few $\mu\text{g/mL}$. In addition, the degree of hypochromism can be a good indication of how closely the aromatic chromophores are packed together. By following the changes in the intensity measurement at 260 nm, one can monitor the course of heat denaturation, the action of a DNase, or the extent of drug binding (especially intercalation), because the base stacking in DNA should decrease as these events proceed. Figure 1.27b shows the result of tracing the change of the 260 nm intensity by raising the temperature of the solution [129]. It is seen that a transition occurs in a narrow temperature range, with a sharp ascent, instead of a gentle slope. This is called a “melting curve” of the DNA, as it refers to the dissociation of the duplex into two separate strands, resulting in an increased disorder in the base stacking. The temperature at the midpoint of the transition is called the *melting temperature*, and denoted by T_m . The melting temperature T_m of DNA in solution varies according to the base composition and the salt concentration. The T_m of natural B-DNA plotted against the GC content forms a straight line (see Fig. 1.28), from ($T_m = 81^\circ\text{C}$, GC = 30%) to

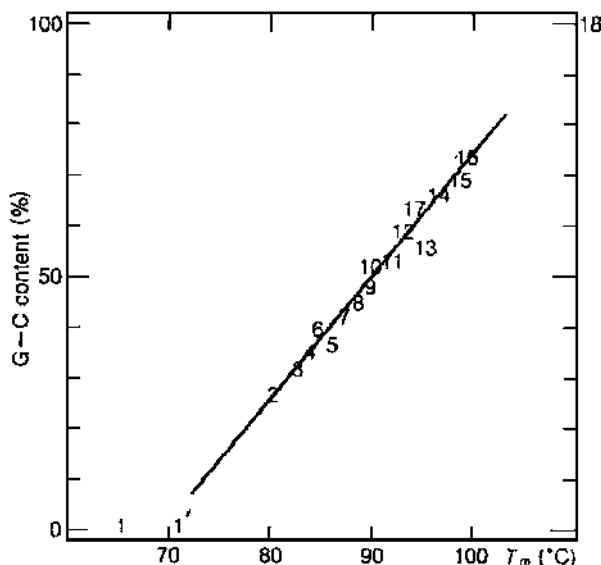


FIGURE 1.28. Relation between the melting temperature T_m of DNA duplex and its guanine + cytosine content (GC content) [63]. Solvent: 0.15 M NaCl + 0.015 M Na citrate. DNA samples are (1) poly(dA-dT)•poly(dA-dT), (1') poly(dA)•poly(dT), (2) DNA from *Clostridium perfringens*, (3) from *Bacillus cereus*, (4) from *B. thurigiensis*, (5) from *Proteus vulgaris*, (6) from phage T6, (7) from *B. subtilis*, (8) from *Vibrio cholerae*, (9) from phage T7, (10) from *E. coli*, (11) from *Neisseria meningitidis*, (12) from *Aerobacter aerogenes*, (13) from *Azotobacter vinelandii*, (14) from *Mycobacterium plei*, (15) from *Sarcina lusia*, (16) from *Micrococcus lysodeikticus*, (17) from *Pseudomonas aeruginosa*, and (18) poly(dG)•poly(dC).

($T_m = 100^\circ\text{C}$, GC = 74%), as long as the salt concentration of the solution is kept constant (0.15 M NaCl plus 0.015 M Na citrate) [129]. The transition moment of the 260 nm band is in the plane of the base, perpendicular to the helical axis in B-DNA, as shown by flow linear dichroism (FLD) experiments [130] (Section 1.2.3).

1.2.2. Fluorescence Spectra

The low-temperature fluorescence and phosphorescence spectra of the common mononucleotides [131] are shown in Fig. 1.29. At room temperature and in neutral aqueous solutions, the fluorescences of the common base residues are very weak, as they have quantum yields of only 3×10^{-6} , 2×10^{-6} , 4×10^{-5} , 3×10^{-5} , and 1×10^{-5} , respectively, for 5'-AMP, 5'-GMP, 5'-TMP, 5'-CMP, and 5'UMP [131]. In contrast, many of the modified and unusual bases involved in natural nucleic acids are strongly fluorescent. Four examples of such fluorescence spectra are shown in Fig. 1.30 [132–136]. Their strong fluorescence is useful in designing biochemical and biophysical experiments of various systems involving these base residues. For example, chloroacetaldehyde can react with the adenine bases in DNA to give the fluorescent product, ethoxyadenosine (Fig. 1.30a). This was used to study the

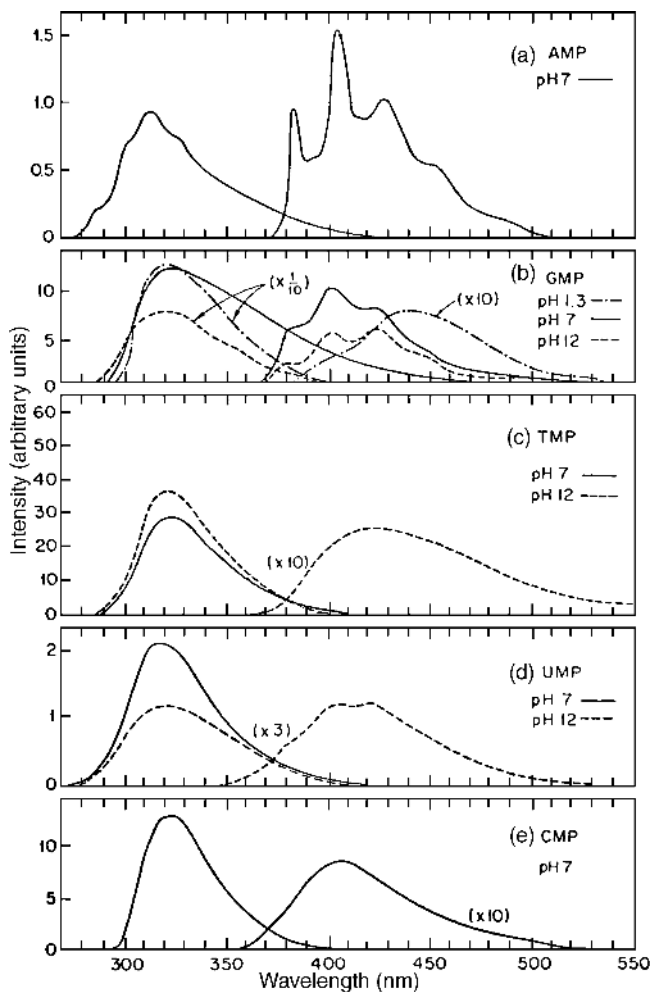


FIGURE 1.29. The fluorescence (near 310 nm) and phosphorescence (near 410 nm) spectra of nucleotides [126]. The solvent was an ethylene glycol : water (1 : 1) glass and spectra were obtained at 80 K.

availability of positions 1 and 6 of adenine in DNA for intermolecular reaction with the helix-destabilizing protein [132].

1.2.3. Linear and Circular Dichroism

1.2.3.1. Linear Dichroism Linear dichroism (LD) measures the differences in the absorption by a molecule when it is irradiated by parallel ($A_{\parallel}$) versus perpendicular ($A_{\perp}$) plane-polarized light. As this technique is very general, arising from asymmetry in the molecular structure or environment, it has found many

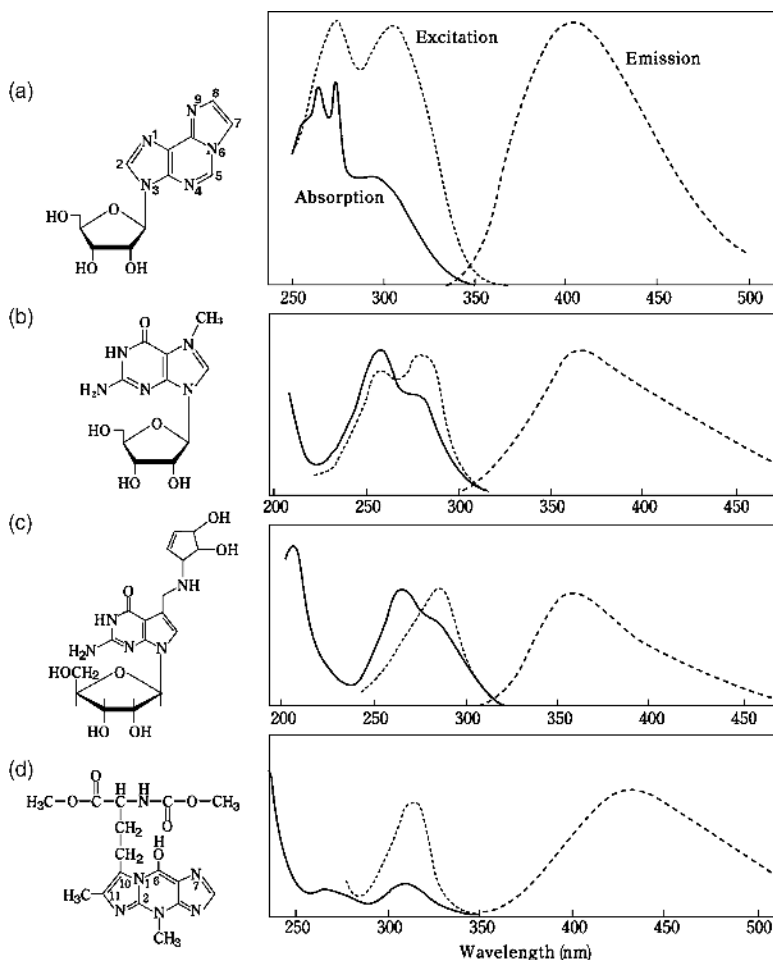


FIGURE 1.30. The fluorescence and fluorescence excitation spectra of four less common, naturally occurring bases of nucleotides [132–136] (—) absorption spectra; (----) emission spectra; (.....) excitation spectra; (a) ethoxyadenosine; (b) 7-methylguanosine in aqueous solution at pH 4.5; (c) Q-base nucleoside-5'-phosphoric acid in aqueous solution at pH 6.2; (d) Yt base in aqueous solution at pH 6.2.

kinds of uses in biomolecular studies [137–139]. Linear dichroism is used most frequently to measure the average relative orientation of a molecule when it is in an aligned matrix, such as when a probe molecule is in a membrane, or when an aromatic molecule is in the presence of the ordered structure of the bases in a DNA duplex.

The measured absorption bands may be in the UV–visible region or in the infrared (IR) region. The former generally provides information based on an electronic–vibrational envelope, whereas the latter probes individual vibrational modes within

the ground electronic state. The general equation for linear dichroism is

$$LD = A_{\parallel} - A_{\perp} \quad (1.3)$$

It is often helpful to express these results as the dichroic ratio, by dividing LD by the absorbance of a randomly aligned sample A_r or $(A_{\parallel} + A_{\perp})$. This ratio allows one to calculate the order parameter S of the sample [140,141].

The sample can be a crystal, deposited as a film on a window, or oriented by an electric field (ELD; see Sections 3.5) or by a rotating flow cell. In the last case, it is called *flow-aligned LD* (FLD). When a sample rotates or flows through a narrow space while the spectrum is measured, the shearing forces have the effect of orienting the molecules along the direction of flow. Using flow to align a sample is common in most other spectroscopic techniques as well. In all cases, the sample is only partially aligned. Only in studies of crystals can full alignment be achieved.

Linear dichroism has been used to observe the transitions between different forms of DNA, as well as the effects of drug binding [142,143]. As stated in Section 1.1.1, the sugar moiety in a nucleotide has a D-stereoisomeric configuration, and must therefore have optical activity. This means that a nucleotide can rotate the plane of polarization of linearly polarized light. With duplex DNA, the helical axis may be aligned along the laboratory z axis (such as by flow), but the bases will be perpendicular to it. Light polarized perpendicular to the helical axis is absorbed by the DNA bases because the electronic transitions of the bases all involve electron movement within their planes. Light polarized parallel to the helical axis is not absorbed. The difference is therefore a negative signal. Single-stranded DNA, however, is too flexible to show any signal. Hence, this technique can distinguish differences in the length and rigidity of polymeric DNA, as each of these affects the hydrodynamic properties of the molecule.

This technique can also be used to study drug-binding to DNA, as a free ligand in solution has no average orientation, and hence no signal. When it is bound to DNA, however, its orientation is fixed with respect to the helical axis. If the molecule is intercalated between its bases, it will have a negative signal, as do the bases (see above). If the molecule is bound in either the major or minor grooves, then it can have a positive signal, depending on the orientation of the chromophore relative to the helix [142].

1.2.3.2. Circular Dichroism Circular dichroism (CD) measures the differences in the absorption of a molecule by left and right circularly polarized light. These differences arise from asymmetry in the structure of the molecule. It is particularly useful for assessing the secondary structure of proteins and nucleic acids, especially the extent to which it is helical. The CD spectrum describes the average molecular conformation in the bulk sample. As such, it cannot determine the orientation of any specific base or residue, but it is very helpful in determining whether a macroscopic structural change occurs when solution or temperature conditions are changed, or when a drug (or ligand) has bound to the biomacromolecule. Note that unlike linear dichroism (LD), CD does not require an oriented sample. For DNA, the spectrum is usually measured between 210 and 290 nm, where the aromatic bases absorb and the spectra are strongly affected by stacking interactions. Unstructured, single-stranded,

DNA and the different DNA helical forms (for duplex, triplex, and quadruplex DNA) can usually be distinguished on the basis of their unique CD profiles [144].

This optical activity can be detected as the difference in absorption intensities for left (ϵ_L) and right (ϵ_R) circularly polarized light [145]. The amount of circular dichroism (CD) is usually given by molar ellipticity (θ)

$$\theta = 3300 \Delta\epsilon \quad (1.4)$$

where $\Delta\epsilon$ is the difference in molar extinction coefficients ($\epsilon_L - \epsilon_R$). It is also possible to perform CD measurements on individual bands within a vibrational spectrum (e.g., an IR spectrum), in which case it is termed *vibrational circular dichroism* (VCD) (Section 6.1.1.3).

The CD spectra of nucleic acids, in general, reflect the relative locations and orientations of the chromophores (i.e., base residues) within the same molecule. As an example, CD spectra of four tetranucleotide sequences are compared in Fig. 1.31

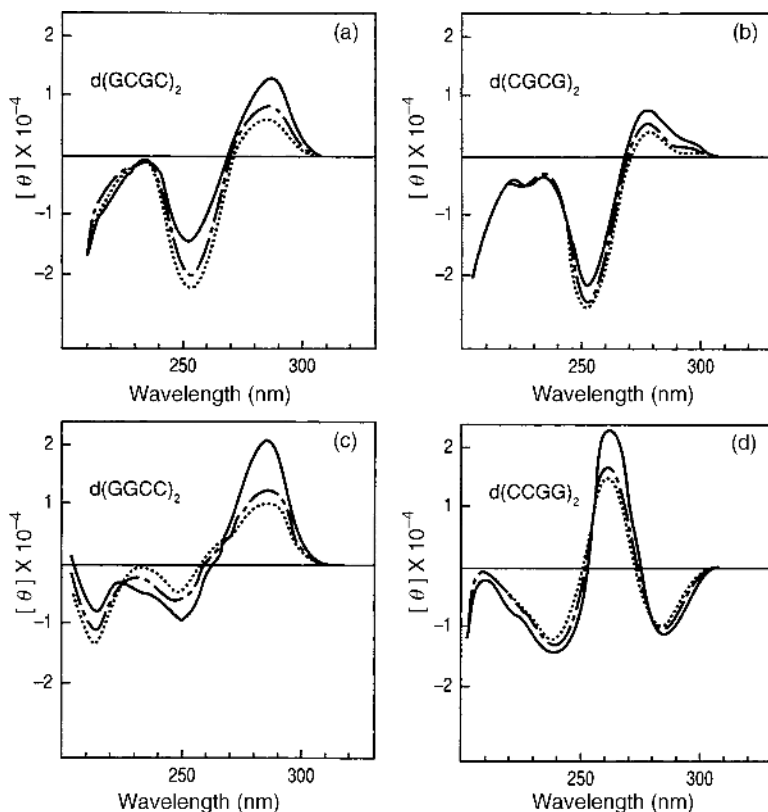


FIGURE 1.31. The circular dichroism (CD) spectra of four tetranucleoside triphosphates with two cytidine and two guanosine residues [146]. Concentrations of the oligomer molecules are all in 130 OD units (absorbance/mole). Solvent: (—), no salt, at 5°C; (---) 0.5 M NaCl, at 5°C; (·····) 0.5 M, at -5°C.

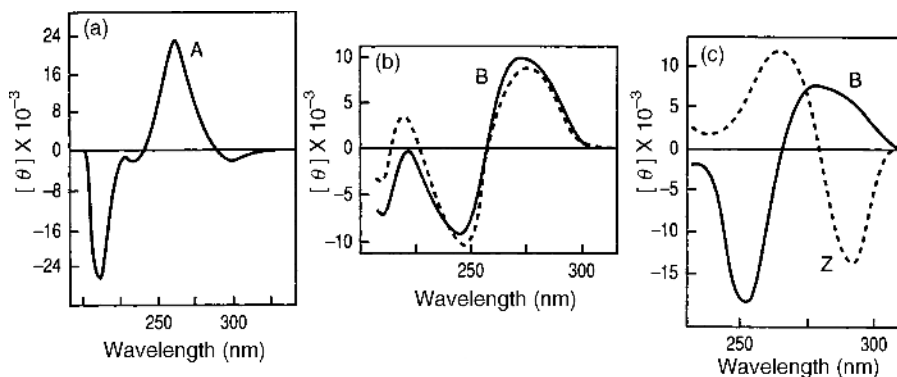


FIGURE 1.32. The circular dichroism (CD) spectra of (a) rice dwarf virus RNA (A-form duplex) [147], (b) (—) DNA from *Micobacterium plei* (B-form duplex) and (---) DNA from *E. coli* (B-form duplex) [148], and (c) poly(dG–dC)•poly(dG–dC) in low ionic-strength solvent (B-form duplex, ---) and in high-ionic-strength solvent (Z-form duplex, —) [149].

[146]. Here, all the molecules have the same number and the same types of chromophores; thus, every molecule has two cytosines and two guanines. Nevertheless, their CD spectra vary significantly. Even more pronounced differences are seen among the A, B, and Z duplexes of polynucleotides, as shown in Fig. 1.32 [147–149].

1.3. VIBRATIONAL SPECTRA

1.3.1. Infrared Spectra of Bases

In general, infrared (IR) absorptions are caused by the vibrations of atoms in a molecule. An IR spectrum is usually presented as a curve having the wavenumber scale (cm^{-1}) along the abscissa and the percent transmittance along the ordinate. It then consists of a number of minima, whose positions correspond to the characteristic vibrational frequencies of the molecule. These characteristic vibrational frequencies correspond to a unique set of ways in which a specific molecular structure can vibrate, and are called “normal modes” of vibration, or often, “normal vibrations.” The number of normal modes for a given molecule is $3N - 6$, where N is the number of atoms in the molecule. As with all absorption processes, the depth of each minimum (or the absorption intensity of each band) is a measure of the amount of dipole oscillation caused by the vibration in question. The vibrational frequencies of a molecule depend on the geometrical structure and the interatomic forces of the molecule [150]. The atomic masses of the chemical groups participating in the vibration are a particularly important determinant of the vibrational frequencies. An easy and advantageous experimental approach is therefore to replace the exchangeable hydrogens (^1H) with the slightly heavier isotope,

deuterium (^2H). This is called *deuteration*, and the resulting difference in mass, although small, can shift the frequency enough to allow determination of the atoms involved in the vibration.

Thymine. Thymine is a molecule consisting of 15 atoms, and therefore it has 39 normal vibrations. Approximate descriptions of these 39 vibrational modes are listed in the first column of Table 1.5, and the lower trace of Fig. 1.33 shows the infrared absorption spectrum of thymine [151,152]. Most of the observed bands assigned to these modes, including their frequencies and relative intensities, are given in the second column. These assignments were based on the results of the following studies:

- (1) Comparison with the Raman spectrum (see discussion below) of thymine.
- (2) Examination of the effect of deuteration, $\text{N1-H} \rightarrow \text{N1-D}$ and $\text{N3-H} \rightarrow \text{N3-D}$, on the IR spectrum.
- (3) Comparison with the results of *ab initio* self-consistent field molecular orbital (SCF-MO) calculations of the thymine molecule for the equilibrium geometry, harmonic force constants, vibrational frequencies, vibrational modes, and IR and Raman intensities [151].
- (4) Comparison with the known correlation of the vibrational modes and the vibrational frequencies of the benzene molecule. These ring vibrational modes [150,153,154] are shown in Fig. 1.34, and their characterizations are given in the upper half of Table 1.6 (under “six-membered ring”).

In the $1750\text{--}1650\text{ cm}^{-1}$ region, two strong infrared bands appear at 1738 and 1675 cm^{-1} (Fig. 1.33, lower trace). The former is a nearly pure C2=O stretching, and the latter is an in-phase stretching of the C4=O and C5=C6 bonds. In the $1500\text{--}1400\text{ cm}^{-1}$ region, *ab initio* calculations predict two stronger bands due to the ν_6^g ring vibration (see Fig. 1.34) and N3-H in-plane bending vibration (modes 14 and 15 of Table 1.5). The latter is expected to be shifted to a lower frequency on deuteration. Actually, two stronger IR absorptions are observed at 1450 and 1430 cm^{-1} for thymine and apparently only one at 1470 cm^{-1} for thymine- d_2 . The stronger band at $\sim 1200\text{ cm}^{-1}$ is assignable to a N3-C4/N1-C6 in-phase stretching mode (mode 18 in Table 1.5). It is generally accepted that the N-H out-of-plane bending vibration of a nucleic acid base gives a broad and strong IR band, whereas the corresponding Raman band is extremely weak. In fact, thymine shows a broad and strong infrared band at 846 cm^{-1} , which is removed by deuteration, whereas practically no Raman band is found in this vicinity (see Fig. 1.33). The theoretical calculations predict that mode 24 (Table 1.5) should produce a strong infrared band at 867 cm^{-1} , and that the corresponding Raman band should be very weak. Therefore, mode 24 is unquestionably assigned as the 846 cm^{-1} band. The calculations also predict the N1-H out-of-plane bending vibration (mode 29) at 645 cm^{-1} , but a strong and broad IR band is actually observed at 559 cm^{-1} (not shown in Fig. 1.33).

Cytosine. The infrared absorption spectrum of cytosine is shown in the lower trace of Fig. 1.35 [155,156]. This molecule has 13 atoms, and therefore 33 normal vibrations.

TABLE 1.5. Vibrational Modes of Thymine Molecule with IR and Raman Bands Assigned to These Modes

Mode Number	Vibrational Mode	Observed Frequencies (cm ⁻¹) and Intensities	
		IR	Raman
1	N1—H stretching	—	—
2	N3—H stretching	—	3304(w) ^a
3	C6 stretching	~3100(w) ^a	3089(w) ^a
4	Methyl deg stretching A'	2963(s) ^a	2966(s) ^a
5	Methyl deg stretching A''	2934(s) ^a	2934(s) ^a
6	Methyl sym stretching	2934(s) ^a	2934(s) ^a
7	C2=O stretching	1738(s)	1700(w)
8	C4=O, C5=C6 in-phase stretching	1675(s)	1672(s)
9	C4=O, C5=C6 out-of-phase stretching	—	1652(w)
10	Methyl deg deformation A'	—	1489(w)
11	N1—H in-plane bending	1480(m)	1459(w)
12	Methyl deg deformation A''	—	—
13	Methyl sym deformation	—	1489(w)
14	v _b ⁶ ring vibration	1450(m)	1434(w)
15	N3—H in-plane bending	1430(m)	1408(w)
16	C6—H in-plane bending	1365(w)	1368(s)
17	Kekulé ring vibration	1242(m)	1246(m)
18	v _a ⁶ ring vibration + C5—Cm stretching	1200(s)	1214(w)
19	v _{δB} ⁶ ring vibration	1156(w)	1156(m)
20	Methyl rocking A'	1030(m)	1048(m)
21	Methyl rocking A''	—	—
22	C6—H out-of-plane bending	985(m)	983(m)
23	v _{δa} ⁶ ring vibration	937(m)	934(w)
24	N3—H out-of-plane bending	—	846(bs)
25	Tr ring deformation + C5—Cm stretching	814(s)	804(m)
26	Ring out-of-plane vibration	809(w)	—
27	Ring out-of-plane deformation	762(m)	—
28	Ring breathing	743(m)	741(s)
29	N1—H out-of-plane bending	700(bm)	—
30	Ring deformation	602(w)	616(s)
31	δ _a ⁶ ring vibration	559(s)	565(m)
32	δ _b ⁶ ring vibration	—	478(m)
33	Ring out-of-plane vibration	—	428(m)
34	C2=O, C4=O in plane bending	—	—
35	Ring out-of-plane vibration	—	320(w)
36	C5—Cm in-plane bending	—	286(w)
37	Ring out-of-plane vibration	—	—
38	Methyl torsion	—	—
39	Ring out-of-plane vibration	—	—

Key: A'—in-plane mode, A''—out-of-plane mode; sym.—symmetric, deg—degenerate; v_a⁶ etc.—see Fig. 1.33; Cm—carbon atom of CH₃ group; s—strong, w—weak, m—medium.

^aObserved for thymidine.

Source: Refs. 151 and 152.

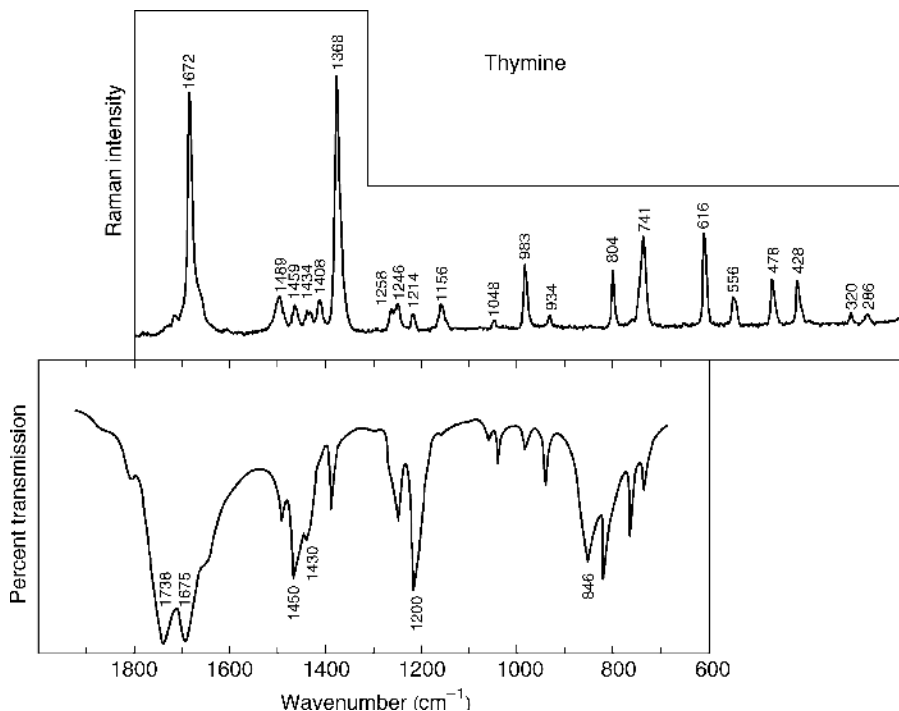


FIGURE 1.33. Infrared spectrum (lower trace) and Raman spectrum of thymine in crystalline powder [151]. The Raman spectrum was excited at 488.0 nm.

Of these, 23 ($= 2N - 3$) are in-plane and 10 ($= N - 3$) are out-of-plane vibrations. Most of the out-of-plane vibrations are weak in both IR and Raman spectra, and their vibrational modes are not always clear. Thus, they are less useful for the study of drug–DNA interactions. Therefore, only the vibrational modes of in-plane vibrations are given in Table 1.7 [156]. The assignments of these modes to the observed IR absorption bands were made in a manner similar to that for thymine, described above. The frequencies and intensities of the observed infrared bands are listed in the second column of Table 1.7.

The strong IR bands at 3380 and 3169 cm^{-1} of cytosine are assigned to the NH_2 antisymmetric and symmetric stretching vibrations, respectively, because they are shifted to 2545 and 2337 cm^{-1} when deuterated (cytosine-1, amino- d_3). It is known that the hydrogen-bonded imino group often gives a broad band at a much lower frequency. The broad IR band of cytosine around 2700 cm^{-1} and that of cytosine-1, amino- d_3 around 2150 cm^{-1} can be assigned to the N1–H and N1–D stretching vibrations, respectively.

In the 1800–850 cm^{-1} region, the highest two observed frequencies, 1703 and 1662 cm^{-1} , are respectively assigned as the C2=O and C5=C6 (or ν_b^6) stretching vibrations. Molecular orbital calculations [156] indicate that the next highest frequency should be that of NH_2 scissoring motion, and the strong IR band at 1615 cm^{-1}

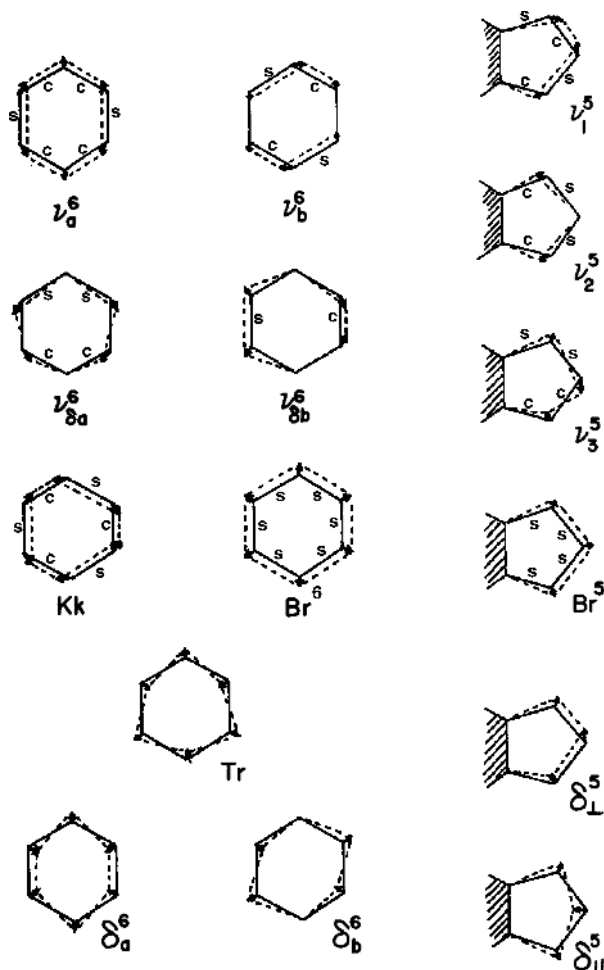


FIGURE 1.34. Standard modes of vibration of six-membered ring, and their abbreviated designation [153,154] (s—stretch; c—contract).

corresponds to it. As expected, it is shifted down to 1182 cm^{-1} on deuteration. The next three highest frequencies should be assigned to ν_a^6 of the hexagon, C4-(NH₂) stretching, and N1-H in-plane bending vibrations that correspond, respectively, to the observed 1538 , 1505 , and 1465 cm^{-1} IR bands. When deuterated, the first two remain nearly unchanged, while the last one (N1-H bending) shifts to 950 cm^{-1} . Molecular orbital calculations predict that the C5-H/C6-H in-phase bending vibration should come next, followed by the Kekulé vibration, and next the C5-H/C6-H 180° out-of-phase bending. Corresponding to these are three prominent IR bands, which are observed at 1364 , 1277 , and 1236 cm^{-1} , respectively. The NH₂ rocking vibration is expected to be the next highest one, and the 1150 cm^{-1} band can be assigned to it.

TABLE 1.6. Designation of Ring Vibrations

Origin of Vibration	Description
Six-membered ring	
ν_a^6 and ν_b^6	$1595\text{ cm}^{-1}(e_{1g})$ of benzene
$\nu_{a\delta}^6$ and $\nu_{b\delta}^6$	$1484\text{ cm}^{-1}(e_{1u})$ of benzene
Kk	Kekulé vibration, $1308\text{ cm}^{-1}(b_{2u})$ of benzene
Br ⁶	Six-membered ring breathing vibration, $992\text{ cm}^{-1}(a_{1g})$ of benzene
Tr	Triangle vibration, $1010\text{ cm}^{-1}(b_{1u})$ of benzene
Da6 and db6	$605\text{ cm}^{-1}(e_{2g})$ of benzene
Five-membered ring (fused with another ring through the left vertical bond)	
ν_1^5	Antisymmetric stretching vibration with respect to twofold axis along horizontal direction
ν_2^5	Symmetric stretching vibration
ν_3^5	Another antisymmetric stretching vibration
Br ⁵	Five-membered ring stretching vibration
$\delta_{ }^5$	Symmetric deformation vibration ^a
$\delta_{\perp}^5$	Antisymmetric deformation vibration ^a

^aIf the left (left-hand-side) vertical bond is fixed, the center of gravity of the five-membered ring moves either parallel (||) or (⊥) to the horizontal line (molecular axis).

Sources: Refs. 150, 153 and 154; see also Fig. 1.34.

Markedly broad and strong infrared bands at 823 and 520 cm^{-1} of cytosine (Fig. 1.35) are not predicted to be in-plane vibrations; they must be N1–H out-of-plane bending and NH₂ wagging vibrations. On deuteration (to cytosine-1,amino-*d*₃), they are shifted to 605 cm^{-1} (N1–D out-of-plane bending) and 388 cm^{-1} (ND₂ wagging), respectively.

Adenine. The infrared absorption spectrum of adenine is shown in the lower trace of Fig. 1.36 [157,158]. The adenine molecule consists of 15 atoms, and hence it should have 39 normal vibrations. Five of these are stretching vibrations involving N–H and C–H bonds, with frequencies higher than 2000 cm^{-1} (not shown in Fig. 1.36). In the region below 1800 cm^{-1} , 34 normal vibrations are expected, and their proposed assignments [158] are listed in the second column of Table 1.8. As indicated in the table, 12 out of these are out-of-plane vibrations, while the remaining 22 are in-plan vibrations. Thirteen (13) of the modes involve mostly hydrogen displacements, while 21 of the modes involve mostly skeletal motions of the purine ring. The modes of the latter category can be correlated to the known modes of benzene and imidazole, which are listed in Table 1.9, with a numbering scheme ranging from ν_1 to ν_{21} .

Guanine. The infrared spectrum of guanine [159] is shown in the lower trace of Fig. 1.37. This molecule has 29 in-plane normal vibrations. Table 1.10 lists the descriptions of these vibrational modes (column 1), along with the frequencies of the observed IR bands that have been assigned to them (column 2) [160].

TABLE 1.7. Vibrational Modes of Cytosine Molecule with the IR and Raman Bands that have been Assigned to These Modes

Mode Number	Vibrational Mode ^a	Observed Frequencies (cm ⁻¹) and Intensities ^b	
		IR	Raman
1	NH ₂ antisymmetric stretching	3380(s)	3354(m)
2	N1–H stretching	2700(m)	—
3	NH ₂ symmetric stretching	3169(s)	3176(m)
4	C5–H stretching	—	3117(m)
5	C6–H stretching	—	3050(w)
6	C=O stretching	1703(w)	1694(m)
7	C5=C6 stretching	1662(s)	1653(m)
8	NH ₂ scissoring	1615(s)	1612(w)
9	ν_a^6	1538(m)	1533(s)
10	C4–N stretching	1505(m)	1498(m)
11	N1–H bending	1465(s)	1462(m)
12	C5–H and C6–H in-phase bending	1364(m)	1361(s)
13	Kekulé vibration	1277(m)	1276(s)
14	C5–H and C6–H out-of-phase bending	1236(m)	1247(s)
15	NH ₂ rocking	1150(w)	1148(w)
16	$\nu_{\delta a}^6$	1100(w)	1108(m)
17	Triangle vibration	966(w)	971(m)
18	$\nu_{\delta b}^6$	—	894(w)
19	Ring breathing vibration	793(w)	792(s)
20	δ_a^6	600(m)	597(s)
21	δ_b^6	549(m)	546(s)
22	C=O in-plane bending	533(m)	533(s)
23	C–N in-plane bending	—	400(m)

^aSee Table 1.5.^bSee Refs. 154–146.

In the frequency region of 4000–2000 cm⁻¹, special care must be taken in correlating the calculated and observed vibrations, because many of the IR bands here are severely perturbed by hydrogen bonding. *Ab initio* molecular orbital (MO) calculations [160] show that the vibration at 3686 cm⁻¹ is the amino antisymmetric stretching and that at 3496 cm⁻¹ is the amino symmetric stretching. In the IR spectrum, on the other hand, 2′/3′-benzyliden-5′-tritylguanosine dissolved in deuterated chloroform shows two bands at 3515 and 3407 cm⁻¹ due to the molecule free from any hydrogen bonding [161]. These two bands may correspond to the two vibrations described above:

In the 1200–800 cm⁻¹ region, the contribution of the C8–H deformation mode is well established [160]. The vibrations calculated to be at 1174 cm⁻¹ for guanine, at 885 cm⁻¹ for guanine-8-*d*, at 1146 cm⁻¹ for guanine-1,9-amino-*d*₄, and at 930 cm⁻¹ for guanine-*d*₅, should involve a large contribution of C8–H or C8–D in-plane

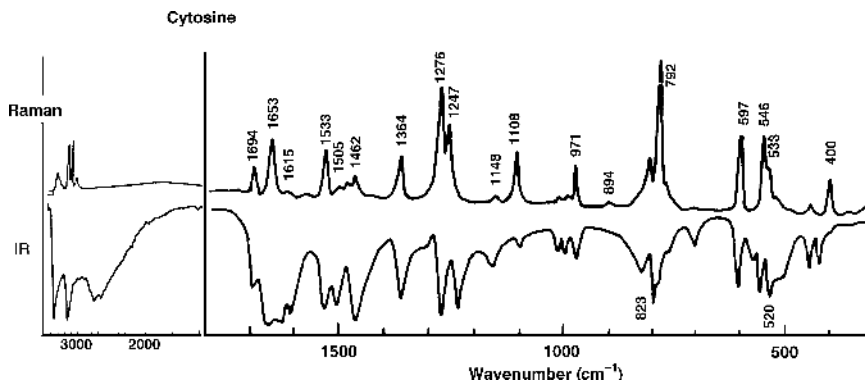


FIGURE 1.35. Raman and infrared spectra of crystalline cytosine (powder) [156].

bending mode. These vibrations are correlated with observed IR bands at 1174, 926, 1168, and 897 cm^{-1} , respectively.

1.3.2. Raman Spectra of Bases

The Raman spectra of thymine [151], cytosine [151], adenine [158], and guanine [162,163] are shown in the upper traces of Figures 1.33, 1.35, 1.36 and 1.37, respectively. Raman spectra are generally obtained by using excitation wavelengths in the visible region. When light having a frequency ν_e strikes a molecule, an electronic polarization is induced within it. This polarization has the same frequency as the incident light (ν_e), and causes light to be scattered with that same frequency, ν_e . This is

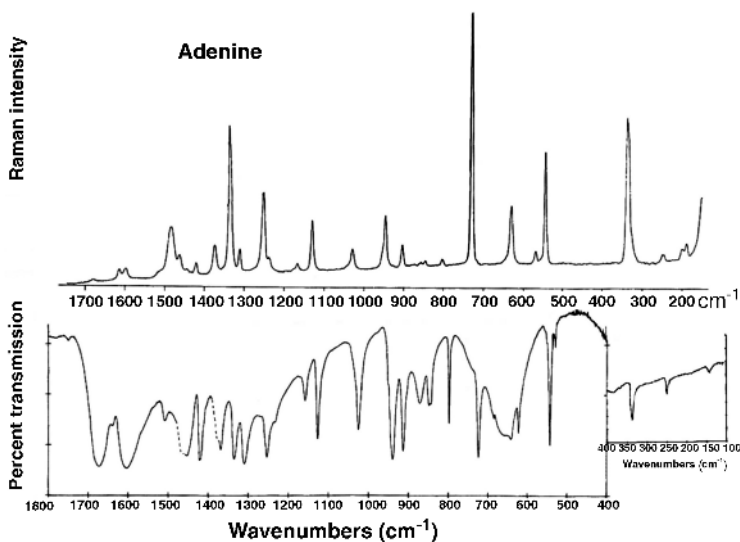


FIGURE 1.36. Vibrational spectra of adenine [158].

TABLE 1.8. Vibrational Modes of Adenine Molecule with IR and Raman Bands Assigned to These Modes

Mode Number	Vibrational Mode ^a	Observed frequencies (cm ⁻¹) and Intensities	
		IR	Raman
1	NH ₂ scissoring	1673.2(s)	1676.6(w)
2	v ₁	1637.6(w)	—
3	v ₂	1604(s)	1621.1(w), 1591.1(w)
4	v ₃	1507.8(m)	1514.0(w), 1507.8(w)
5	v ₄	—	1485.3(m), 1481.6(m)
6	N9—H bending	1451.4(m), 1441.8(w)	—
7	v ₅	1421.1(s)	1418.9(w)
8	C2—H bending	1368.0(m)	1370.8(m)
9	v ₆	1334.7(s)	1332.3(s)
10	v ₇	1308.7(s)	1307.0(m)
11	C6—NH ₂ stretching	1252.8(s)	1248.7(m)
12	v ₈	1234.0(w)	1234.8(w)
13	C8—H bending	1156.8(m)	1163.3(w)
14	v ₉	1125.9(s)	1126.0(m)
15	NH ₂ rocking	1024.7(s)	1024.4(w)
16	v ₁₄	951(w)	852(w)
17	C8—H out-of-plane bending	940.3(s)	941.4(m)
18	v ₁₁	913.3(s)	898.8(w)
19	C9—H out-of-plane bending	872.1(m)	—
20	C2—H out-of-plane bending	849.2(m), 844.0(m)	845.5(w), 839.9(w)
21	v ₁₆	797.1(m)	797.0(w)
22	v ₁₀ (ring breathing)	723.3(s)	723.6(s)
23	v ₁₉	684.2(w)	—
24	NH ₂ wagging	660(s)	—
25	v ₂₀	640.8(m)	—
26	v ₁₂	621.6(m)	622.8(m)
27	v ₁₇	—	560.2(w)
28	v ₁₃	543.0(s)	535.1(m)
29	v ₁₅	—	530(w)
30	v ₁₈	529.9(w)	—
31	NH ₂ torsion	380(w)	—
32	C6—NH ₂ out-of-plane bending	337.1(m)	—
33	C6—NH ₂ in-plane bending	—	331.3(s), 327.2(s)
34	v ₂₁	294.3(m)	242.0(w), 238.3(w)

^aSee Table 1.5.

Source: Ref. 158.

called *Rayleigh scattering*. In addition, when a molecule vibrates, its polarizability oscillates with the same frequency as the vibration (ν). All of these processes occur simultaneously. Therefore, when light with a frequency ν_e strikes a vibrating molecule, the light is scattered as a linear combination of the frequencies, $\nu_e \pm \nu$. This is called *Raman scattering*. A *Raman spectrum* is a plot of the scattering intensity as a

TABLE 1.9. Classification of the 21 Ring Vibrations of Adenine with Respect to Origin

Type of Motion: Bond Stretching or Angle Bending	Origin		Mode Number
	Six-Membered Ring	Five-Membered Ring	
In-plane stretching	1595 cm ⁻¹ (e_{2g}) of benzene		ν_1, ν_2
	1485 cm ⁻¹ (e_{1u}) of benzene		ν_3, ν_4
		Double-bond stretching, 1446 cm ⁻¹ of imidazole	ν_5
	Kekulé vibration, 1380 cm ⁻¹ (b_{2u}) of benzene		ν_6
		1324 cm ⁻¹ of imidazole	ν_7
		Single-bond stretching 1261 cm ⁻¹ of imidazole	ν_8
	Ring breathing, 992 cm ⁻¹ (a_{1g}) of benzene		ν_9
		1142 cm ⁻¹ of imidazole	ν_{10}
In-plane bending	Triangle vibration		
	1010 cm ⁻¹ (b_{1u}) of benzene		ν_{11}
	605 cm ⁻¹ (e_{2g}) of benzene		ν_{12}, ν_{13}
		Elongation to the long axis of the molecule	ν_{14}
		Rocking $\perp$ to the long axis of the molecule	ν_{15}
Out-of-plane bending	Crown vibration		
	703 cm ⁻¹ (b_{2g}) of benzene		ν_{16}
	405 cm ⁻¹ (e_{2u}) of benzene		ν_{17}, ν_{18}
		Twisting 619 cm ⁻¹ of imidazole	ν_{19}
		Envelope formation, 660 cm ⁻¹ of imidazole	ν_{20}
		Butterfly motion	ν_{21}

Source: Ref. 158

function of frequency, ν . The frequency of a Raman band corresponds to a vibrational frequency ν of the molecule, and its intensity is proportional to the square of the amplitude of the scattered light. Since a molecule vibrates in more than one way, as given by its normal modes, a Raman spectrum is composed of a complex mixture of all of those vibrations, ν .

The frequency of a Raman band, in general, is expected to be the same as that of an IR band for the same molecular vibration. In the case of thymine, all of the prominent

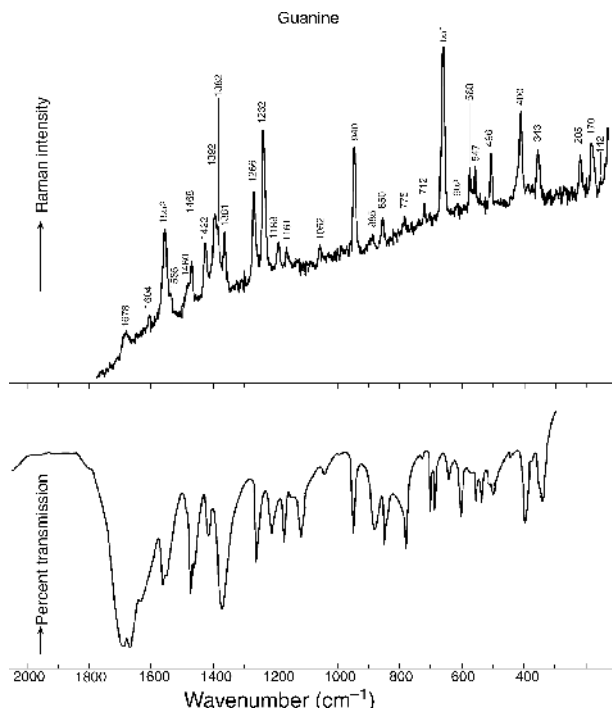


FIGURE 1.37. Vibrational spectra of guanine [159,162].

Raman bands have counterparts of nearly the same frequencies in the IR spectrum, and vice versa (see Fig. 1.33), however, the frequencies of each counterpart are not always exactly the same (Table 1.5). In this example, these differences arise because the sample used for these IR and Raman investigations was a crystalline powder of thymine, not an isolated thymine molecule. In the crystalline environment, some intermolecular vibrational couplings take place (space group $P2_1/c$), causing each molecular vibration to split into a group of vibrations with slightly different frequencies. Thus, the selection rules are different for Raman scattering and IR absorption in the crystalline state.

The relative intensities of Raman versus IR bands are seldom the same, as they arise from different selection rules. For example, the carbonyl ($C=O$) stretching vibrations always give strong IR bands (1738 cm^{-1} for thymine, 1703 cm^{-1} for cytosine, and 1702 cm^{-1} for guanine), whereas the corresponding Raman bands are weak. Ring-breathing vibrations always exhibit strong Raman bands ($741, 792, 724$, and 651 cm^{-1} for thymine, cytosine, adenine, and guanine, respectively), whereas IR counterparts are weak. In drug–DNA interaction studies, a difference in the intensity of a Raman band compared to its IR counterpart can be an interesting and useful indicator of binding.

As with other optical spectroscopic techniques, Raman spectroscopy can often yield useful information when the polarization of the signal is analyzed. This is

TABLE 1.10. Vibrational Modes of Guanine Molecule

Mode Number	Vibrational Mode ^a	Observed Frequency (cm ⁻¹)	
		IR	Raman
1	NH ₂ antisymmetric stretching	3515	—
2	N9—H stretching	—	—
3	N1—H stretching	—	—
4	NH ₂ symmetric stretching	3403	—
5	C8—H stretching	3125	—
6	C6=O stretching	1702	1678
7	C2—N stretching + NH ₂ scissoring	1675	—
8	ν_a^6 ring vibration	1638	1604
9	NH ₂ scissoring	1565	1552
10	ν_b^6 ring vibration	1550	1535
11	Kekulé + ν_1^5	1477	1480
12	Kekulé - ν_1^5	1464	1468
13	$\nu_{\delta\beta}^6$ ring vibration	1420	1422
14	ν_2^5 ring vibration	1375	1392
15	N1—H in-plane bending	1362	1361
16	N9—H in-plane bending	1261	1266
17	Br ⁵	1174	1188
18	C8—H in-plane bending	1216	1232
19	ν_3^5	1150	1161
20	NH ₂ rocking	1118	—
21	$\nu_{\delta a}^6$	1042	1052
22	$\delta_{ }^5$	950	940
23	Tr + δ_{+}^5	850	850
24	$\delta_b^6 + \delta_{\perp}^5$	703	712
25	Br ⁶ + Br ⁵	645	651
26	δ_b^6	557	568
27	δ_a^6	501	496
28	C2—N in-plane bending	404	400
29	C6=O in-plane bending	348	343

^aSee Table 1.5 for ring vibrations.

Source: Ref. 160.

because the Raman scattering light is partially polarized even in the case of randomly oriented molecules in a solution sample. In such isotropic samples, the components of the induced electric dipole change as a function of the spatial coordinates of the molecule. The signal intensity is thus measured both parallel and perpendicular to the plane of the polarization of the incident light. This is called the *Raman depolarization ratio*:

$$\rho = \frac{I_{\perp}}{I_{||}} \quad (1.5)$$

When an isotropic sample (e.g., a solution) is excited with plane-polarized light, the totally symmetric vibrations will produce Raman bands that are strongly polarized in the plane parallel ($I_{\parallel}$) to that of the incident light. In this case, ρ is small or zero, and the band is referred to as being *polarized*. Nontotally symmetric vibrations, however, produce bands which are called *depolarized*. These have a $\rho \approx 0.75$, meaning that the intensity perpendicular ($I_{\perp}$) to the incident light is 75% as strong as from the parallel orientation ($I_{\parallel}$).

The effects of polarized excitation is more complex when the sample is in the crystalline state, where its orientation is fixed with respect to the incident light and detector. Here, the components of the polarization depend on the orientation of the crystal axes with respect to the direction of the incident light, its plane of the polarization, and the placement and polarization of the detector. This requires the measurement of the different components of the polarized Raman intensities (I_{aa} , I_{bb} , I_{cc} , etc.) along the crystal axes (i.e., a, b, and c), where the orientations of incident light and the detector are indicated. More details on depolarization studies may be found in Appendix A.1. A more quantitative discussion requires discussion of the polarizability tensor of an ellipsoid [164].

1.3.3. Vibrational Spectra of Nucleosides

Most of the molecular vibrations of a nucleoside are localized primarily in either the base residue or the sugar moiety. Thus, the Raman spectrum of thymidine shows Raman bands attributable to both the thymine residue and the deoxyribose moiety. The origins of these bands have been determined by examining the Raman spectrum of thymidine-1',2',3',4',5'-(^{13}C)₅, in which the five carbon atoms in the deoxyribose moiety were uniformly replaced with the ^{13}C isotope [165]. This heavier isotope causes the vibrational frequencies within the sugar component of thymidine- $^{13}\text{C}_5$ to shift with respect to those of natural thymidine. Consequently, those Raman bands that shift may be attributed to deoxyribose motions. The medium-intensity band at 898 cm^{-1} showed the greatest frequency shift, -17 cm^{-1} . It is assigned to a ring-breathing vibration of the deoxyribose residue [165].

To understand the vibrational modes of the C2' methylene group in deoxyribonucleosides, the thymidine stereoisotopomer, (2'S)-[2'- ^2H]thymidine, was synthesized to incorporate deuterium in the S configuration at the furanosyl 2' carbon [166]. The crystalline powder of this molecule was analyzed by IR and Raman spectroscopies (the latter using excitation at 1063 nm and 514.5 nm) and were compared with those of normal thymidine. On the basis of these experiments, detailed vibrational characterizations were made for many of its observed Raman and IR bands (shown in part in Table 1.11). For example, the strong Raman band at 898 cm^{-1} , which was assigned to a deoxyribose ring-breathing vibration (see above), turns out to be sensitive to C2'(S) deuteration. Thus, this band arises from a vibration that also has a significant amount of 2'/CH₂ motion, probably methylene rocking. However, a vibration that consists entirely of a methylene rocking motion is expected to exhibit a ^{13}C shift larger than 20 cm^{-1} . Shifts of that magnitude were not observed in this region of the thymidine spectrum, indicating that thymidine has no purely methylene rocking mode. Instead,

TABLE 1.11. Vibrational Bands of the C2'-Methylene Group of Thymidine

Bands of C2'H ₂ Species			Bands of C2'H(S)D Species		
Infrared ^a	Raman ^a	Assignment ^b	Infrared ^a	Raman ^a	Assignment ^b
2994(m)	2996(s)	CH ₂ antisym st	2984(w)	2984(s)	C—H st
2956(w)	2956(m)	CH ₂ sym st	2187(w)	2187(m)	C—D st
1410(w)	1404(w)	CH ₂ scissor	1078(s)	C2'DH def	
1173(s)	1174(m)	CH ₂ wag	990(s)	993(m)	C2'DH def
1069(s)	1103(w)	CH ₂ twist	903(m)	905(m)	C2'DH def
896(w)	898(s)	CH ₂ rock	799(m)	801(m)	C2'DH def

^aBand frequencies are given in wavenumber units (cm⁻¹), and relative intensities are indicated in parentheses.

^bKey: sym—symmetric, antisym—antisymmetric, st—stretch, def—deformation, scissor—scissoring, wag—wagging, twist—twisting, rock—rocking.

Source: Ref. 166.

the 2'CH₂ rocking modes are likely coupled with deoxyribose skeletal modes (such as breathing motions), and may contribute to several bands in the 700–1000 cm⁻¹ region. It is suggested, however, that the 2'CH₂ contribution is especially large for the 898 cm⁻¹ vibration [166]. By synthesizing [5'-¹³C]thymidine, which has the C5' position selectively replaced with ¹³C, it is possible to identify those bands in the Raman spectrum that arise from motions involving that atom [167]. Whereas most of the Raman bands remained the same as in natural thymidine, several were shifted significantly to lower frequencies. These are listed in Table 1.12, along with their assigned vibrational modes.

A comparison of the Raman spectra of thymine and thymidine reveals that 22 of the Raman bands from the nucleoside correlate directly with those of the base [167]. These are listed in Table 1.13 (for the assignments based on the thymine vibrational analysis, see Table 1.5). Similar frequency correlations can be tabulated for the IR spectra of

TABLE 1.12. Raman Bands Showing Prominent Frequency Shifts on Replacing 5'-¹²C by ¹³C in the Thymidine Molecule

Frequency (cm ⁻¹)		Shift	Vibrational Mode
5'- ¹² C Species	5'- ¹³ C Species		
2934	2931	-3	5'CH ₂ symmetric stretching
1481	1477	-4	5'CH ₂ scissoring
1066	1060	-6	5'C—4'C—H bending
1029	1022	-7	5'CH ₂ rocking
1016	1012	-4	O—5'C—4'C antisymmetric stretching
852	849	-3	O—5'C—4'C symmetric stretching
396	393	-3	O—5'C—4'C bending

Source: Ref. 167.

TABLE 1.13. Raman Bands Common to Thymidine and Thymine

Thymidine		Thymine	
Frequency (cm ⁻¹)	Intensity ^a	Frequency (cm ⁻¹)	Intensity ^a
1692	w	1700	w
1665	s	1672	s
1643	m	1652	w
1483	m	1489	m
1458	w	1459	w
1437	w	1434	w
1390	w	1408	m
1365	s	1368	s
1233, 1227	s	1258, 1246	m
1199	s	1214	m
1124	w	1156	m
1067	m	1048	w
1016	m	983	m
974	w	934	w
794	m	804	m
772	s	741	s
675	s	616	s
565	s	556	m
495	s	478	m
379, 397	m	320	w
297, 305	m	286	w

^aRelative Raman intensity (s—strong, m—medium, w—weak) in the spectra of crystalline powder.

Source: Ref. 167.

thymidine and thymine. However, the IR intensities are more variable, presumably owing to the greater sensitivity of infrared transition moments to the effects of intermolecular interactions, which differ significantly in thymidine and thymine crystal lattices. In the Raman spectra, the observed differences in frequency and intensity between thymidine and thymine are attributable mostly to the base–sugar vibrational couplings. Thymine exhibits, for example, a strong band at 616 cm⁻¹ that has been assigned [151] to the in-plane bending mode in which the six-membered ring undergoes a skeletal deformation, while the C2=O and C4=O groups simultaneously bend in phase with one another (a “windshield-wiper” motion). In thymidine, this vibration is upshifted to 675 cm⁻¹ in the Raman spectrum, possibly due to coupling with a deoxyribose mode at 632 cm⁻¹. In thymidine-1',2',3',4',5'-¹³C₅, this mode is observed at 671 cm⁻¹ in the Raman spectrum [165]. For thymidine-C2'-*d*, the corresponding mode occurs at 663 cm⁻¹ in the Raman spectrum [166]. The C2'-deuteration shift of 12 cm⁻¹ (675 → 663 cm⁻¹) in the Raman spectrum demonstrates vibrational coupling with the deoxyribose moiety.

The strategy of site-specific furanose deuteration was initiated by Harada and co-workers [168], who investigated the coupling of the base and sugar vibrations in purine

ribonucleisides. These workers synthesized and obtained vibrational spectra of adenosine and guanosine that are deuterated at the C1' carbon, the site of attachment of the purine base.

Vibrational coupling between the base and sugar in a nucleoside varies according to the sugar conformation. For example, the ring-puckering vibration of guanine in guanosine produces a Raman band at $682 \pm 3 \text{ cm}^{-1}$ if the ribose has *S*(C2'-*endo*)-*anti* conformation, at $671 \pm 2 \text{ cm}^{-1}$ for *S*(C2'-*endo*)-*syn*, at $664 \pm 2 \text{ cm}^{-1}$ for *N*(C3'-*endo*)-*anti*, and at $625 \pm 2 \text{ cm}^{-1}$ for *N*(C3'-*endo*)-*syn* conformation. Such an empirical rule was derived from a series of Raman spectra of crystals containing the guanosine residues, whose structures are known through crystallographic analyses [169]. Other marker bands were also proposed (Section 6.1.1).

1.3.4. Vibrational Spectra of Nucleotides

In the vibrational spectra of nucleotides and polynucleotides, there appear some Raman and IR bands that are assigned to vibrations localized in their phosphate portions, in addition to those assigned to their nucleoside moieties. These bands, along with their properties, are listed in Table 1.14 [170,171]. Mononucleotides and terminal residues of polynucleotides have a phosphomonoester group. At pH higher than 6, this group takes the form of $-\text{O}-\text{PO}_3^{2-}$, which shows a strong and broad IR band at 1100 cm^{-1} and a strong and sharp band at 980 cm^{-1} in the IR spectrum. The Raman band corresponding to the latter is strong and polarized (see discussion below). The former, at 1100 cm^{-1} , is assigned to the PO_3^{2-} degenerate stretching vibration and the latter at 980 cm^{-1} to the PO_3^{2-} symmetric stretching vibration. These two characteristic bands of the PO_3^{2-} group are seen in the spectra of adenosine-5'-monophosphate, shown in Fig. 1.38 [171,172].

As examples of nucleotide Raman spectra in solution, those of 5'-CMP and 5'-dAMP are shown in Fig. 1.39 panels (a) and (b), respectively [173]. Two scans are seen in each of these, which were measured such that the polarization of the scattered light is selected according to whether it is parallel ($\parallel$) or perpendicular ($\perp$) to

TABLE 1.14. Characteristic Frequencies of the $>\text{PO}_2^-$ and $-\text{PO}_3^{2-}$ Groups

Group	Frequency (cm^{-1})	Infrared Band	Transition Moment	Raman Band	Assignment
$>\text{PO}_2^-$	1230	Strong	Along the O . . . O	Weak	PO_2^- antisymmetric stretching
	1090	Strong	Along the bisector of $\angle \text{OPO}$	Strong	PO_2^- symmetric stretching
$-\text{PO}_3^{2-}$	1100	Strong, broad	$\perp$ to the threefold symmetry axis	Weak	PO_3^{2-} degenerate stretching
	980	Strong, sharp	$\parallel$ to the threefold symmetry axis	Strong	PO_3^{2-} symmetric stretching

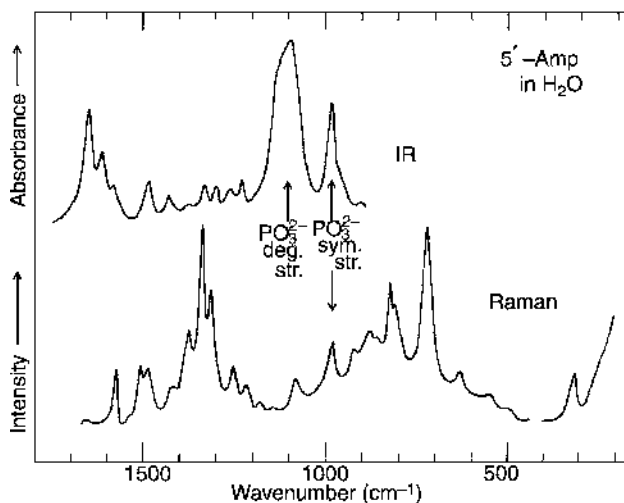


FIGURE 1.38. Infrared and Raman spectra of 5'-AMP in H_2O solution [159,160].

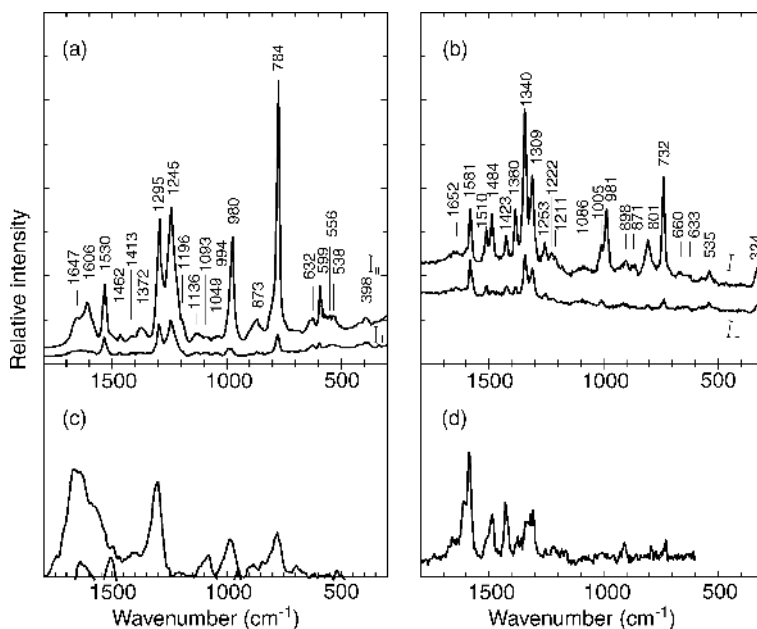


FIGURE 1.39. Raman spectra ($1800\text{--}300\text{ cm}^{-1}$) of some nucleotides in H_2O solutions: (a) polarized Raman spectra of 5'-CMP excited at 488.0 nm ($I_{||}$ —intensity of Raman scattering beam polarized in same direction as that of exciting beam polarization; $I_{\perp}$ —intensity of Raman scattering beam polarized in direction perpendicular to that of exciting beam polarization [173]) of; (b) polarized Raman spectra of 5'-dAMP excited at 488.0 nm . [$I_{||}$ and $I_{\perp}$ are of same width as in (a) [173]]; (c) UV resonance Raman spectrum of 5'-CMP excited at 200 nm [174]; (d) UV resonance Raman spectrum of 5'-dAMP excited at 200 nm [175].

the incident excitation. Together, they constitute the “depolarization ratio” of each Raman band, as discussed Section 1.3.2 and Appendix A.1.

1.3.5. Resonance Raman Spectra

Resonance Raman spectroscopy is essentially the same technique as Raman spectroscopy, except that the energy of the incident light is tuned to an absorbance maximum of the molecule in the UV–visible region. Under this condition, the input energy is the same as (or, is “in resonance with”) the energy gap between the electronic ground and excited states that are associated with the absorbance feature. When the energy of excitation is properly tuned, it will specifically excite vibrations associated with that excited state, resulting in a significant enhancement of the Raman signal. An excited electronic state is a real energy state of the molecule that has energy higher than that of the ground state. This excited state has an electron distribution different from that of the ground state, and often, a somewhat different geometric structure. This allows the Raman transitions to satisfy certain special selection rules. This is in contrast to standard, or “nonresonant” Raman spectroscopy, in which the excitation energy corresponds to that of “*virtual* excited states,” which consist of a large number of overlapping rotation–vibration energy levels within the electronic ground state.

Since each pyrimidine and purine base of nucleic acids has a few strong electronic absorption bands in the 200–280 nm range (Section 1.2.1), the resonance Raman spectroscopy in this wavelength region serves as a powerful tool for the conformational study of nucleic acids with a high sensitivity and selectivity. With the 200–280 nm excitation, the Raman scattering intensity is, in general, so high that the concentration of a nucleotide solution as low as 10^{-4} M is sufficient for observing Raman band and determining its frequency. In addition, the relative intensities of the Raman bands of the same molecule may change considerably as the excitation wavelength of the incident light is tuned to different absorption bands. It is thus often possible to choose the excitation wavelength to maximize the intensity of a particular Raman band of interest relative to other bands. This effectively increases the selectivity of the technique for conformational studies.

To illustrate the marked changes of the absolute intensities of the Raman bands of nucleotides, resonance Raman spectra of 5'-CMP [174] and 5'-dAMP [175] are shown in Fig. 1.39 panels (c) and (d), respectively, along with the off-resonance Raman spectra (Fig. 1.39a, b). When the Raman scattering intensity is plotted as a function the excitation wavelength, it is called an *excitation profile*. Excitation profiles were reported for many of the bands of 5'-CMP (1650, 1529, 1294, 1243, and 784 cm^{-1} bands). Similar excitation profiles were obtained for some Raman bands of 5'-AMP and 5'-dAMP. The intensities of most of the AMP Raman bands are maximized at 260 nm, where the absorption maximum of AMP is located.

1.3.6. Vibrations of the Phosphodiester Linkage

Nucleic acids have an intense and polarized Raman band between 1000 and 1100 cm^{-1} , which can be assigned to the symmetric P–O stretching mode

within the phosphodioxy (PO_2^-) group. This can be done on the basis of detailed analyses of the Raman and infrared spectra of orthophosphates and their simple alkyl esters [176,177]. Likewise, intense bands in the $750\text{--}850\text{ cm}^{-1}$ region can be assigned to vibrations that involve a large contribution from phosphorus–oxygen stretching of the phosphodiester linkages. As there is significant vibrational coupling between the sugar and phosphate groups, these bands are extraordinarily sensitive to the nucleic acid backbone conformation and can provide a firm basis for DNA and RNA conformational analysis. Figure 1.40 shows the Raman frequencies and depolarization ratios of dimethylphosphate, together with descriptions of the assigned vibrational modes, on the basis of a preliminary mathematical analysis.

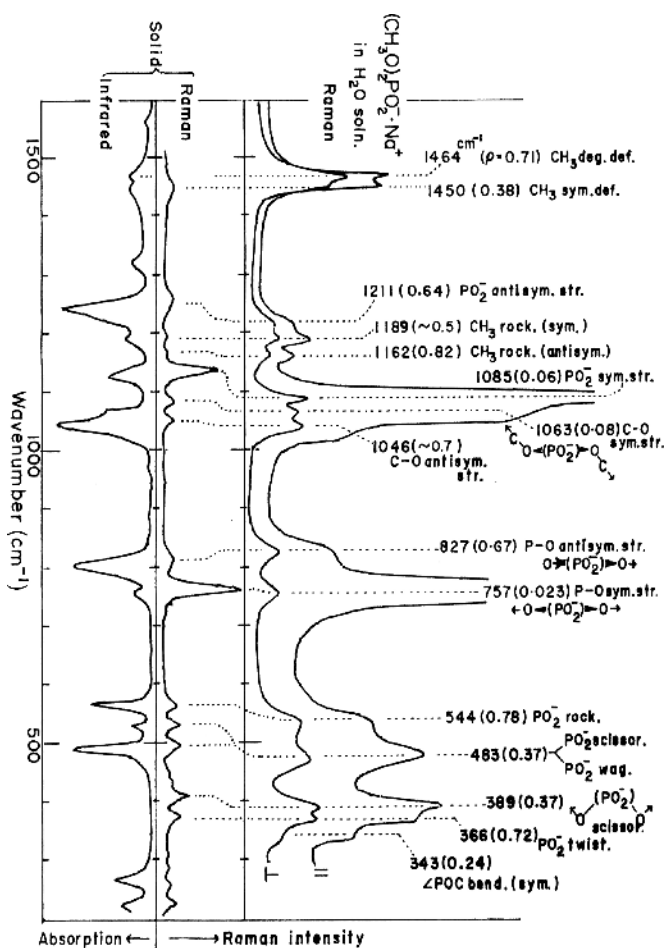


FIGURE 1.40. Raman and infrared spectra, Raman depolarization ratios, and preliminary vibrational assignments of dimethylphosphate anion [177].

Note that the symmetry of the dimethylphosphate anion leads to distinct symmetric (757 cm^{-1}) and antisymmetric (827 cm^{-1}) stretching modes with the expected relative intensities: The 757 cm^{-1} vibration is strong in Raman and weak in IR, whereas 827 cm^{-1} is weak in Raman and strong in IR spectra. However, in nucleic acids, the strict twofold local symmetry is lost in the 3',5'-diester form, and it is not possible to rigorously define the symmetric and antisymmetric character of the phosphodiester O—P—O stretching bands. This may account for the phosphodiester vibrations of DNA and RNA (approximately $810\text{--}840\text{ cm}^{-1}$) exhibiting both relatively high frequencies and relatively high Raman (and also IR) intensities.

1.3.7. Infrared Spectra of DNA

Infrared absorption spectra of typical DNA, as undeuterated, partially deuterated, and deuterated films, are shown in Fig. 1.41 [178]. Since the film was kept at 92% relative humidity, the DNA molecules are considered to be B-DNA. As seen in the figure, B-DNA gives a broad and strong band at 3400 cm^{-1} (or 2500 cm^{-1}), which is assigned to the O—H (or O—D) stretching vibration of the hydrated water molecules. The bands in the $1800\text{--}1500\text{ cm}^{-1}$ region are assigned to the C=O stretching, skeletal stretching, and N—H bending vibrations of the base residues [137]. They are partially overlapped by HOH (or HOD) bending vibrations of the water molecules. A strong band at 1220 cm^{-1} is assigned to the PO_2 antisymmetric stretching vibration, and another strong band at 1100 cm^{-1} is assignal to the PO_2 symmetric stretching vibration of the phosphodiester group [137,170]. There are a number of bands with strong, medium, and weak intensities in the region between 1000 and 700 cm^{-1} . These bands are due to the P—O stretching, C—O stretching, and N—H out-of-plane bending vibrations and also to liberations (incomplete rotation in a film or crystalline state) of adsorbed water molecules. The N—D out-of-plane bending liberation of D_2O should appear in the $600\text{--}400\text{ cm}^{-1}$ region. Assignments of some of these bands are indicated in Fig. 1.41.

Figure 1.42 shows the infrared spectra of the A and B forms of an oriented DNA film from *Micrococcus lysodeikticus* observed with polarized radiation [179]. Prominent spectral differences between the A and B forms of DNA are observed in the $1300\text{--}1000\text{ cm}^{-1}$ region [179,180]. In A-DNA, the PO_2^- antisymmetric stretching band (located at 1230 cm^{-1} in parallel-polarized spectrum, and at 1243 cm^{-1} in perpendicular-polarized spectrum) shows a perpendicular dichroism (i.e., $I_{\perp} > I_{\parallel}$), whereas in B-DNA the antisymmetric stretching band (at 1220 cm^{-1} in $\parallel$, and at 1230 cm^{-1} in $\perp$) is almost nondichroic (i.e., $I_{\perp} \approx I_{\parallel}$).

In A-DNA, the PO_2^- symmetric stretching band at 1092 cm^{-1} (both $\parallel$ and $\perp$ spectra) shows a parallel dichroism (i.e., $I_{\perp} < I_{\parallel}$), whereas in B-DNA the symmetric stretching band at 1088 cm^{-1} (both $\parallel$ and $\perp$ spectra) shows a perpendicular dichroism (i.e., $I_{\perp} > I_{\parallel}$). These differences in the IR dichroism are ascribed to the difference in the orientations of the PO_2^- group between A-DNA and B-DNA. Thus, in A-DNA the $\text{O} \cdots \cdots \text{O}$ line of the PO_2^- group is directed nearly perpendicular to the helical axis

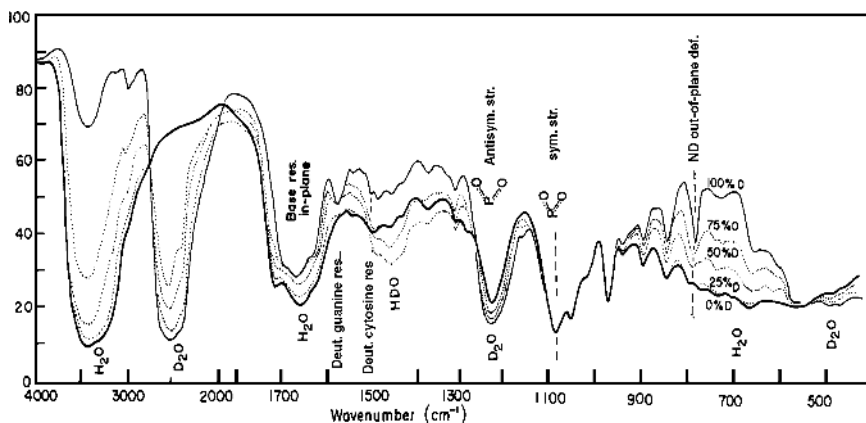


FIGURE 1.41. Infrared absorption spectra of calf thymus DNA film placed in air of 92% relative humidity [178]. The spectra were observed in undeuterated, partially deuterated, and nearly completely deuterated states. Deuteration is shown on the right side of the figure.

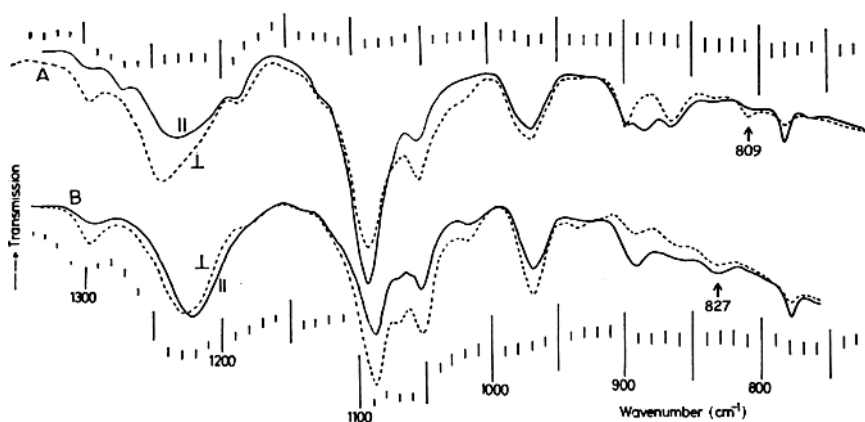


FIGURE 1.42. Infrared absorption curves of oriented films of DNA from *Micrococcus lysodeikticus* in its A and B forms, observed with polarized radiation [179]. The solid line represents the spectrum when the electric vector of the incident radiation is parallel to the fiber axis; the broken line, the spectrum when the electric vector is perpendicular to the fiber axis. The film was prepared by evaporating an aqueous solution of DNA with 10^{-5} M NaCl onto an AgCl plate. The spectrum of the A form was obtained by maintaining a prolonged environmental humidity of 75%. The B-form spectrum was easily obtained at 92% humidity. The conversion from A to B occurs readily when the humidity is changed from 75% to 92%, but the conversion of B form back to the A form in the film requires a constant environmental 75% humidity lasting for a week or so (Section 1.1.7).

TABLE 1.15. DNA Samples Used in Spectroscopic Studies of Fig. 1.43

Source	CG Content (%)
<i>Tetrahymena pyriformis</i> GL	25 ^a
<i>Clostridium perfringens</i>	31 ^b
Calf thymus	43 ^b
<i>Escherichia coli</i>	51 ^b
<i>Pseudomonas aeruginosa</i>	65 ^a
<i>Micrococcus lysodeikticus</i>	72 ^b

^aEstimated from the melting temperature (61.8°C in 0.015 M NaCl + 0.0015 M Na citrate) and buoyant density in CsCl density gradient centrifugation (1.686, as given by Flavell and Jones [182]).

^bUlizur [183].

^cSueoka and Cheng [184].

Source: Ref. 184.

and the bisector of its O–P–O angle is directed nearly parallel to the helical axis. In B-DNA, on the other hand, the O . . . O line is inclined by around 54° from the helical axis and the bisector is nearly perpendicular to the helical axis.

In the 800–700 cm^{−1} region, A- and B-DNAs show a difference in the vibrational frequency that probably involves mainly the P–O single bond stretching mode (Fig. 1.40). Thus, in A-DNA a weak band is observed at 807 cm^{−1}, while in B-DNA a broad and weak band appears at 832 cm^{−1}. The former shows a perpendicular dichroism, whereas the latter is almost nondichroic. As will be described below, these two diagnostic frequencies, 807 cm^{−1} for A-DNA and 832 cm^{−1} for B-DNA, are observed more conspicuously in Raman spectra.

Nishimura et al. [181] examined the IR spectra (within 1750–1480 cm^{−1}) in D₂O of six different DNA samples, each with different guanine–cytosine (GC) content (Table 1.15) [182–184]. As may be seen in Fig. 1.43, the absorption peaks at 1621, 1642, 1668, and 1694 cm^{−1} become more prominent with decreasing GC content, while those at 1580, 1648, and 1682 cm^{−1} become more prominent as the GC content increases. In fact, after a proper adjustment of the absorbance scale in each absorption curve, these frequencies are found to have an almost linear intensity relationship with the GC content. This result indicates that, on the average, the A-T base pairs (or the G-C base pairs) in all of these DNA samples are in similar structural environments with respect to adjacent base pairs. The vertical stacking interactions between the base pairs in a double-helix polynucleotide structure should certainly affect the IR absorptions of the base pairs, but this was not clearly detected in this series of DNAs, even though their GC content spans the range of 25–72%.

1.3.8. Raman Spectra of B-DNA

Figure 1.44 shows the polarized Raman spectra of a B-DNA fiber obtained with a laser excitation of 514.5 nm [185], and the polarized Raman intensities I_{cc} , I_{bb} , I_{bc} , and I_{cb} were measured in the (cc), (bb), (bc), and (cb) configurations (see Section 1.3.2). Table 1.16 lists frequencies of the Raman bands observed for B-DNA fiber and for B-DNA in solution. The vibrational modes assigned to these bands and depolarization

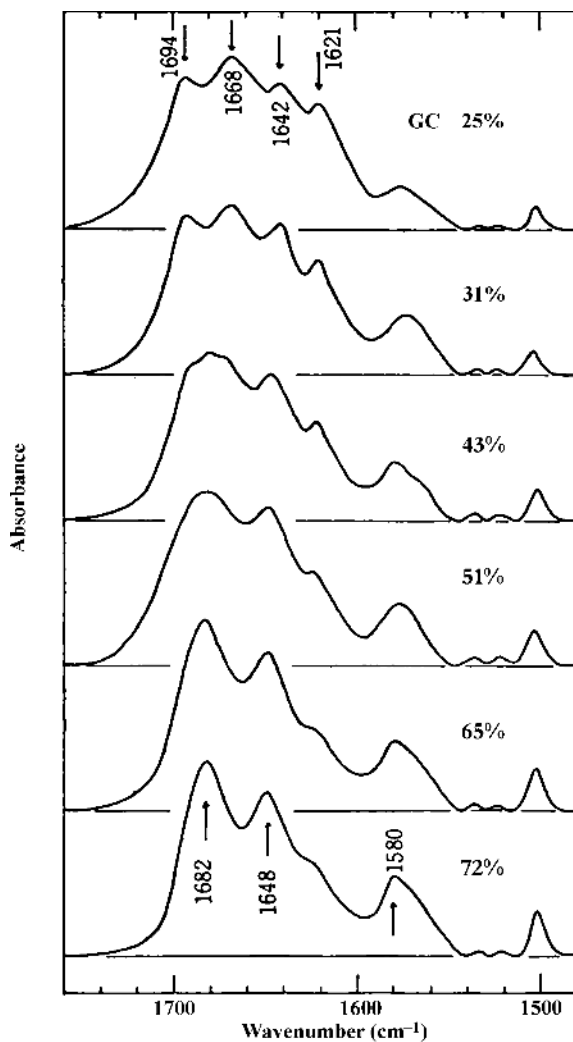


FIGURE 1.43. Infrared absorption curves of different DNA samples in D₂O solutions at about 40°C. Each sample had different GC content (see Table 1.15) [181].

ratios observed for the B-DNA in solution (with random, isotropic orientation) are also given [185]. The 1670 cm^{-1} band of B-DNA is assigned to the exocyclic carbonyl (C=O) bond stretching vibrations of the bases. Among the three base residues containing carbonyl groups, the thymine contribution to the band intensity is large, whereas that of guanine is small, and that of cytosine is negligible. The depolarization ratio ($\rho = 0.26$) observed for the 1670 cm^{-1} band of randomly oriented B-DNA can be explained as a weighted average of the ρ values of the contributing base residues [173], in accordance with their respective intensities.

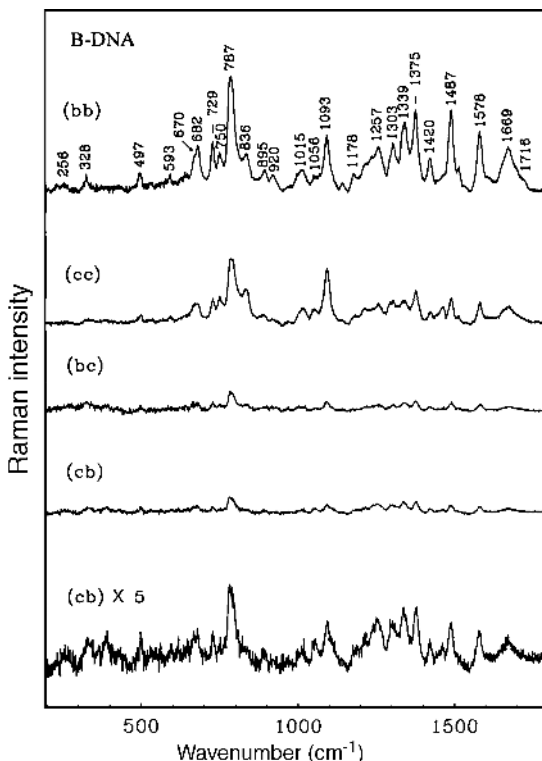


FIGURE 1.44. The Raman spectra (I_{bb} , I_{cc} , I_{bc} , and I_{cb}) of a B-DNA fiber corresponding to four tensor components (bb, cc, bc, and cb, respectively) using 514.5 nm excitation, and at 10°C and 92% humidity [185]. Here a, b, and c refer to the crystal axes. The DNA fibers were about 200 μm thick and were mounted in a thermostatically controlled hygostatic chamber (micro-sample cell), which was designed specifically for the sample platform of a Raman microscope. This allowed the humidity to be carefully controlled, and the sample, well-equilibrated. The electric vector of the exciting radiation was directed along the laboratory x axis. A polarizer located before the entrance slit of the monochromator selects for Raman scattering along either the x or y axis that is perpendicular to it. The fiber axis of the sample was placed along either the x or y axis.

The 1578 cm^{-1} band of B-DNA comprises overlapping contributions from adenine and guanine. The adenine contributor is almost completely depolarized ($\rho = 0.64$), whereas that of guanine is significantly polarized ($\rho = 0.34$) [173]. The intermediate value of ρ (0.54) observed for the 1578 cm^{-1} band of B-DNA is thus consistent with values observed for purine mononucleotides. The 1489 cm^{-1} band of B-DNA, also a composite of contributions from two purines, is very strongly polarized in the nucleic acid ($\rho = 0.17$), as well as in the mononucleotides ($\rho = 0.11$ in GMP and 0.13 in dAMP); however, in this case the purine mononucleotide depolarizations are apparently not simply additive. Interestingly, the depolarization characteristics of the two purine ring modes at 1578 and 1489 cm^{-1} for isotropically oriented DNA in solution are dissimilar.

TABLE 1.16. Raman Bands of B-DNA in Fiber and in Solution

Frequency		Principal Assignment _a	Depolarization Ratio (ρ)
Fiber (cm ⁻¹)	Solution (cm ⁻¹)		
1669	1670	C=O stretching (T)	0.26 ± 0.05
1578	1578	Ring mode (A, G)	0.54 ± 0.05
1487	1489	Ring mode (G, A)	0.17 ± 0.02
1465	—	5'CH ₂ scissoring	—
1420	1420	2'CH ₂ scissoring	0.42 ± 0.02
—	1375	Ring mode (T)	0.27 ± 0.03
1339	1340	Ring mode (G, A)	0.43 ± 0.02
—	1302	Ring mode (A, C)	0.33 ± 0.05
1257	1254	Ring mode (C, T)	0.47 ± 0.05
1093	1092	PO ₂ ⁻ symmetric stretching	<0.04
920	923	Deoxyribose	0.11 ± 0.05
895	900	Deoxyribose	0.36 ± 0.05
836	832	O—P—O stretching	0.17 ± 0.10
792	792	O—P—O stretching	0.13 ± 0.05
784	782	Ring breathing (C)	0.17 ± 0.05
750	749	Ring breathing (T)	0.10 ± 0.05
729	728	Ring breathing (A)	0.13 ± 0.05
682	683	Ring breathing (G)	0.19 ± 0.05
—	670	Ring mode (T)	0.18 ± 0.05
497	497	PO ₂ ⁻ scissoring	0.15 ± 0.10

^aKey: T—thymine; A—adenine; G—guanine; C—cytosine.

Source: Ref. 185.

Depolarization ratios of CH₂ scissoring modes near 1420 cm⁻¹ have been examined in Raman spectra of nucleotides [173]. In every case the ρ value is near 0.4. This is also the case for isotropically oriented B-DNA (see Table 1.16). The four intense Raman bands of B-DNA at 1254, 1302, 1340, and 1375 cm⁻¹ can be assigned to in-plane ring vibrations of the base residues. A common characteristic of each of these Raman bands for isotropically oriented B-DNA in solution is that they will have a depolarization ratio with an intermediate magnitude ($0.25 < \rho < 0.50$). In accordance with the discussion in Section 1.3.2, such a ρ value is indicative of a relatively anisotropic Raman tensor. (Using the notation of Appendix A.1 and Fig. 1A.1, these would be given as $|r_1|, |r_2| > 1$.) Accordingly, these bands are potentially good candidates for probing DNA base orientations in complex biological assemblies by polarized Raman spectroscopy.

The PO₂⁻ symmetric stretching band at 1092 cm⁻¹ is highly polarized ($\rho < 0.04$), and the Raman tensor should be relatively spherical, or isotropic. (Using the notation of Appendix A.1, and Fig. A.1, we expect $r_1 \approx r_2 \approx 1$, i.e., $\alpha_{xx} \approx \alpha_{yy} \approx \alpha_{zz}$.) Accordingly, this band in the polarized Raman spectra is not expected to be a useful indicator of the residue orientation.

Raman bands at 895 and 920 cm^{-1} of B-DNA fiber and those at 900 and 923 cm^{-1} of B-DNA solution can be assigned to bond stretching vibrations of deoxyribose ring. Because of its highly Raman depolarization ($\rho = 0.11$), the 923 cm^{-1} band is considered to represent a more symmetric furanose ring stretching vibration than does the 900 cm^{-1} band ($\rho = 0.36$). These assignments are consistent with those proposed for the sugar group of thymidine [165].

Depolarization ratios for phosphodiester modes of the 750–850 cm^{-1} interval are shown in Table 1.16. It is seen that the bands at 792 cm^{-1} ($\rho = 0.13$) and 832 cm^{-1} ($\rho = 0.17$) are both highly polarized. In B-DNA, therefore, the depolarization ratios suggest that a distinction between symmetric and antisymmetric motions of the C–O–P–O–C bond network is inappropriate (see Section 1.1.8 and Fig. 1.39).

Of the bands assigned to ring-breathing-type motions of the base residues, 683 (G), 728 (A), 749 (T), and 782 (C) cm^{-1} , those of G and C exhibit significantly higher ρ values, that is, lower polarizations, in B-DNA than in the corresponding mononucleotide [173]. This may result from Raman tensor perturbations specific to pairing and stacking of these residues in B-DNA.

1.3.9. Comparison of Raman Spectra of A- and B-DNA

Figure 1.45 shows the polarized Raman spectra of A-DNA fibers [185]. These were obtained through exactly the same procedure with that used for the polarized Raman spectra of B-DNA (Section 1.3.9) except that the relative humidity in the sample chamber was maintained at 75%, instead of at 92%, which is the requirement for B-form DNA.

Raman bands originating from C–C and C–N bond stretching vibrations of the base residues are expected to dominate the 1100–1700 cm^{-1} spectral region. Normal coordinate calculations suggest that many of the modes here involve 180° phase differences in the stretching of adjacent bonds of the heterocycles [153,154]. For both A- and B-DNA fibers, seven strong peaks are detected in this region of the Raman spectrum. In the I_{bb} spectrum of the B-DNA fiber (Fig. 1.44, trace) the peaks are centered at 1257, 1304, 1339, 1375, 1487, 1578, and 1669 cm^{-1} ; the corresponding bands of A DNA (Fig. 1.45, trace) occur at 1257, 1304, 1339, 1376, 1486, 1577, and 1677 cm^{-1} . In B-DNA, the band near 1339 cm^{-1} , due to adenine, is less intense than the thymine band near 1376 cm^{-1} . The relative intensities of these adenine and thymine markers are reversed in A-DNA. Figures 1.44 and 1.45 show that all of the strong I_{bb} Raman bands noted above are greatly reduced in I_{cc} spectra of both A-DNA and B-DNA. This fact may indicate that, for each of the corresponding normal modes of vibration, the out-of-plane tensor component (α_{zz}) is much smaller than at least one of the in-plane tensor components (α_{xx} or α_{yy}) (see Section 1.3.5).

Although in both A and B fibers the base planes are oriented close to the normal to the fiber axis (c axis), the base tilt angle in A-DNA (about 20°) is greater than that in B-DNA (nearly 0°) (Sections 1.1.3 and 1.1.7). Accordingly, it might be anticipated that Raman anisotropies for the above-noted in-plane base vibrations, as measured by the polarized Raman intensity ratio I_{bb}/I_{cc} , should be greater for B-DNA than for A-DNA. However, Figures 1.44 and 1.45 indicate that this is not the case for any of the bands in

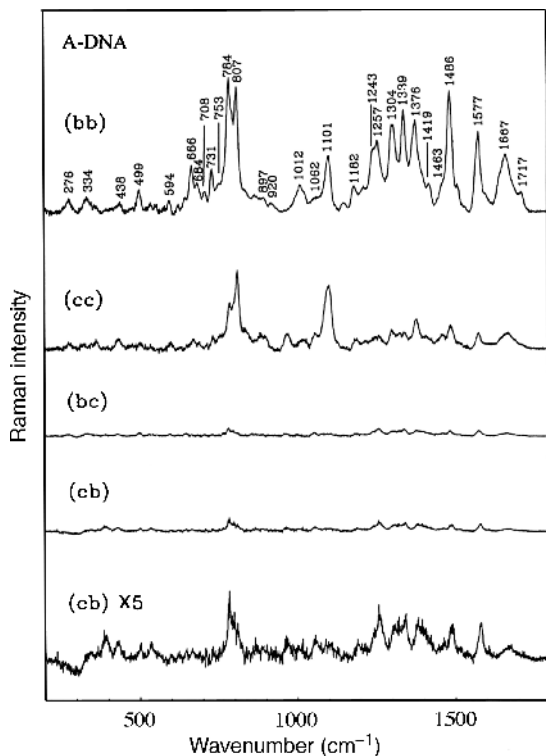


FIGURE 1.45. Raman spectra (I_{bb} , I_{cc} , I_{bc} , and I_{cb}) of an A-DNA fiber corresponding to four tensor components (dd, cc, bc, and cb, respectively). Here a, b, and c refer to the crystal axes. The spectra were acquired by exciting the sample at 514.5 nm and maintaining the temperature at 10°C and the humidity at 75% [185].

question. A possible cause for this is the higher crystallinity and greater degree of unidirectional orientation achievable for fibers of the A form than for fibers of the B form. This is consistent with fiber X-ray diffraction patterns reported for sodium salts of A- and B-DNA [12,13].

The I_{cc} spectrum of A-DNA reveals a very strong Raman band at 807 cm^{-1} for the sugar in the C3'-endo geometry (Fig. 1.45). B-DNA exhibits a much weaker band at 836 cm^{-1} for sugars in the C2'-endo geometry (Fig. 1.44). The 807 and 836 cm^{-1} frequencies correspond to the well-characterized “marker” bands of DNA, which are considered diagnostic of A and B conformations, respectively [177,187,188] (see Section 1.3.10). Extensive study of nucleic acid model compounds [176,177] indicates that these modes originate from vibrations associated with the 3',5'-phosphodiester bond network (C—O—P—O—C) and not from sugar puckering (see also Section 1.3.9).

For B-DNA, an additional marker band near 792 cm^{-1} has been proposed on the basis of deuteration effects observed in fiber of calf thymus DNA [188]. A putative B backbone marker near 790 cm^{-1} would normally be masked by DNA of mixed-base

composition due to the presence of overlapping Raman band near 784 cm^{-1} assignable to a ring mode of the cytosine residue. We see, from the I_{bb} and I_{cc} spectra of A-DNA (Fig. 1.45), that the out-of-plane Raman tensor component (α_{zz}) of the 784 cm^{-1} cytosine mode is relatively small. Accordingly, the strong band in the I_{cc} spectrum of B-DNA cannot be assigned to the cytosine mode. A reasonable assignment is to a phosphorus–oxygen single bond stretching vibration, which is expected in the $750\text{--}850\text{ cm}^{-1}$ region [176].

The two relatively weak bands of B-DNA at 895 and 920 cm^{-1} are assigned to bond stretching vibrations of the deoxyribose ring. This assignment is supported by the observation that similar bands, at 852 and 898 cm^{-1} , in the Raman spectrum of thymidine are shifted by 12 and 17 cm^{-1} , respectively, on ^{13}C substitution of $\text{C1}'$, $\text{C2}'$, $\text{C3}'$, $\text{C4}'$, and $\text{C5}'$ positions of deoxyribose residue [165]. Further support is provided by significant band shifts of guanosine in the $850\text{--}1000\text{ cm}^{-1}$ region on deuterium substitution of the $\text{C1}'$ position [168]. Similarly, weak bands are observed in A-DNA at 882 and 897 cm^{-1} . For the band of 882 cm^{-1} in A-DNA, $I_{bb} < I_{cc}$ is observed. For all other deoxyribose bands in this region, we find $I_{bb} > I_{cc}$.

The Raman band assigned to the symmetric $\text{O}=\text{P}=\text{O}$ stretching vibration of the nucleic acid phosphodioxy (PO_2^-) group is expected near 1093 cm^{-1} in the structures of B form and near 1100 cm^{-1} in A-form structures [137,170]. These PO_2^- markers are observed at the expected frequencies both B and A fibers (see Fig. 1.44 and 1.45). For both DNA conformations, we note also that the I_{bb} and I_{cc} intensities are nearly equal. This contrasts sharply with the results from polarized infrared spectroscopy, where $I_{||} > I_{\perp}$ for A DNA, but $I_{||} < I_{\perp}$ for B-DNA (see Section 1.3.10). The fact that $I_{bb} \approx I_{cc}$ for the Raman PO_2^- mode can therefore be attributed to an intrinsically isotropic local Raman tensor.

The antisymmetric $\text{O}=\text{P}=\text{O}$ stretching vibration of the phosphodioxy group is expected near 1200 cm^{-1} . Although this vibration generates intense infrared absorption, the Raman scattering is characteristically very weak. Actually, no band definitely assignable to the antisymmetric $\text{O}=\text{P}=\text{O}$ stretching mode can be detected in spectra of either the A- and B-DNA fiber (see Figs. 1.44 and 1.45).

Relatively intense Raman bands that can be assigned to the “ring breathing” motions of guanine, adenine, thymine, and cytosine are observed, respectively, at 666 , 731 , 753 , and 784 cm^{-1} in A-DNA and 682 , 729 , 750 , and 787 cm^{-1} in B-DNA. As noted above, the 787 cm^{-1} band of B-DNA actually represents the unresolved center of two overlapping bands due to cytosine (near 790 cm^{-1}) and the B backbone (near 790 cm^{-1}). For all of the ring modes, $I_{bb} > I_{cc}$, which is consistent with the fact that the out-of-plane tensor component (α_{zz}) is smaller than those of in-plane components (α_{xx} and α_{yy}). It should be noted that the Raman anisotropies for the in-plane base modes are greater in A-DNA than in B-DNA.

In both A and B fibers, a weak but prominent band is observed near 498 cm^{-1} . Assignment of this band to a valence angle bending mode (scissoring) of the phosphodioxy group is supported by normal coordinate analysis [189]. High Raman anisotropy ($I_{bb} > I_{cc}$) distinguishes this phosphate group bending vibration from the abovementioned phosphorus–oxygen stretching modes in the $750\text{--}1100\text{ cm}^{-1}$ interval.

All Raman bands in the I_{bc} and I_{cb} spectra (Figs. 1.44 and 1.45) are weak, and therefore their signal-to-noise ratios are low. Nevertheless, some interesting conclusions can be drawn from these observations:

- (1) The results show that $I_{bc} = I_{cb}$ for a given DNA fiber; that is, the Raman spectrum of a fiber oriented along the c axis is invariant to an orthogonal interchange of the polarization directions of the incident and scattered light. This is consistent with precise optical alignment of the Raman microscope-spectrometer system, with the absence of polarization artifacts, and with a high degree of unidirectional orientation of DNA along the fiber axis.
- (2) Comparisons between A- and B-DNAs indicate important tensor similarities and differences. For example, the bands near 1487 , 1578 , and 1669 cm^{-1} exhibit nearly equal intensities and band shapes in A and B fibers. On the other hand, the CH_2 scissoring mode near 1420 cm^{-1} gives a sharp band in B-DNA, but a more diffuse band in A-DNA. The bands at 1303 , 1339 , and 1375 cm^{-1} are prominent only in B-DNA, whereas the 1257 cm^{-1} band is prominent only in A-DNA.
- (3) Finally, of particular interest is the observation that the PO_2^- symmetric stretching band appears much more prominently in the I_{bc} (or I_{cb}) spectrum of B fiber than that of the A fiber. Thus, the different orientations of phosphodioxo groups of A- and B-DNA, which are not obvious in the I_{bb}/I_{cc} ratio, are clearly revealed in the I_{bb}/I_{bc} anisotropy.

1.3.10. Raman Spectra of Z-DNA

As stated in Section 1.1.9, the self-complementary hexamer DNA sequence, $\text{d}(\text{CGCGCG})_2$, forms a left-handed helix (Z-DNA) in the crystalline state, as determined by X-ray diffraction [51,52]. Raman measurements on this crystal are shown in Fig. 1.46. For each intense band in the $300\text{--}1700\text{ cm}^{-1}$ region of the Raman spectrum, the relative scattering intensities, I_{aa} , I_{bb} , and I_{cc} , corresponding to the aa , bb , and cc components of the crystal Raman tensors, were determined [190]. In the crystal, the helical axis of the molecule is aligned nearly along the crystallographic c axis, and the base planes are in the crystallographic ab plane, which is exactly perpendicular to the c axis (see Fig. 1.9c).

As seen in Fig. 1.46, Raman bands of the I_{cc} spectrum at 1579 (G), 1486 (G), 1318 (G), 1264 (C), 784 (C), 670 (G), and 625 (G) cm^{-1} are much weaker than their counterparts in the I_{aa} and I_{bb} spectra. This observation is consistent with the crystal structure, which shows all base-planes arranged nearly in the ab plane, that is, nearly perpendicular to the c -axis. Thus, at least one in-plane component (α_{xx} or α_{yy}) must be much greater than the out-of-plane component (α_{zz}) for the Raman tensor corresponding to each band. The two ring breathing bands of guanine at 670 and 625 cm^{-1} are well known as deoxyguanosine conformation markers (see Section 1.3.2 and Fig. 1.37). The 670 cm^{-1} band is assigned to two terminal dG residues with $\text{C2}'\text{-endo/syn}$ conformation, and the 625 cm^{-1} band is assigned to the four internal dG residues with $\text{C3}'\text{-endo/syn}$ conformation.

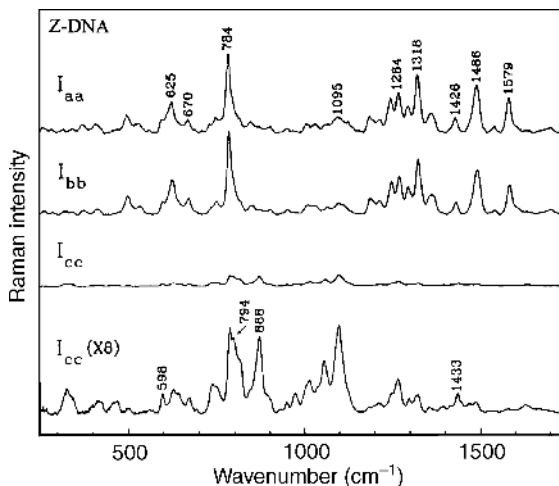


FIGURE 1.46. Raman spectra of the single crystal of $d(CGCGCG)_2$, which has a Z-DNA conformation. The three spectra correspond to the three different orientations of the unit cell axes with respect to the electric vector of exciting and scattered radiation [190].

The weak band at 1426 cm^{-1} in I_{aa} and I_{bb} spectra, assigned to CH_2 scissoring, is not evident in I_{cc} spectrum, where instead a weak band is observed at 1433 cm^{-1} . The crystal structure reveals that most $\text{C2}'\text{H}_2$ groups are oriented such that a line connecting the two hydrogen atoms is approximately perpendicular to the c axis. Since the $\text{H} \dots \text{H}$ line is the direction along which the Raman tensor should have its greatest component [172], the corresponding Raman band should exhibit much greater intensity in the I_{aa} and I_{bb} spectra than in the I_{cc} spectrum. On this basis, the 1426 cm^{-1} band may be reasonably assigned to the $\text{C2}'\text{H}_2$ groups, and the 1433 cm^{-1} band to the $5'\text{CH}_2$ groups.

The 1095 cm^{-1} band, assignable to the symmetric stretching vibration of the anionic phosphodioxy group $(\text{O}=\text{P}=\text{O})^-$, gives similar intensities in all polarized spectra. This is expected because the 10 PO_2^- groups per asymmetric unit of the duplex are oriented along markedly different directions. Additionally, the Raman tensor of this vibration is expected to be intrinsically isotropic.

The Raman band at 868 cm^{-1} draws special attention. This is the only band that is more intense in the I_{cc} spectrum than in either the I_{aa} or I_{bb} spectrum. The Raman tensor of the 868 cm^{-1} band, therefore, exhibits its greatest component along the crystallographic c axis. This band is tentatively assigned to a furanose ring-breathing vibration (see Section 1.3.3 and Ref. [165]). In the $d(CGCGCG)_2$ crystal, all furanose rings are arranged with their “planes” parallel to the c axis and with the $\text{C1}'\text{—O—C4}'$ angle bisectors nearly parallel to the c axis. It is possible that one component of the Raman tensor is localized along the bisector, and that this Raman tensor component is much greater than the other two, thus accounting for the observed anisotropy.

1.4. NMR SPECTRA

Nuclear magnetic resonance (NMR) spectroscopy is a powerful method for studying the structures of drug–DNA interactions, as well as drug-binding affinities. It can be used in different ways to obtain many different types of information; hence, it can be regarded as a family of closely related methodologies, rather than as just one technique. As the details of NMR far exceed the scope of this book, we will consider only the most essential aspects of this subject, focusing primarily on its application to the most common elements in biochemical research, namely, H, C, N, and P. More details on selected aspects of NMR can be found in Appendix A.2. The interested reader is referred to the many excellent texts for information on the principles of NMR [191–196] and on its uses in nucleic acid research [197–203].

The term *nuclear magnetic resonance* (NMR) is derived from its reliance on the intrinsic quantum mechanical properties of the nuclei of atoms. Atoms contain the subatomic particles, electrons, neutrons, and protons, each of which has a property called “spin.” NMR is based on the combined spin of the protons and neutrons in the atomic nucleus, which is described by I , the *nuclear spin*, or *nuclear angular momentum*. By contrast, the related technique, electron spin resonance (ESR), is based on the spin of electrons (see Section 1.5).

Both electrons and nuclei can be regarded as very small bar magnets, with a small *magnetic moment*, μ . The magnitude of μ is dependent on a constant, called the *gyromagnetic ratio*, or γ . Whereas all electrons are identical, the nuclei are not. The values of the nuclear spin and γ value depend on the atomic and mass numbers of a given isotope (see Table 1.17) [204]. Those isotopes having nuclei with large absolute values of γ are easier to detect, and are said to be “sensitive,” whereas those with small γ values are “insensitive.”

The most commonly used isotopes in biological NMR studies are ^1H , ^{13}C , ^{15}N , and ^{31}P (see Table 1.18). Of these, ^1H is the easiest to detect. This isotope is almost 100% abundant and is NMR active—its nucleus has a spin quantum number of $I = \frac{1}{2}$, as well as a large γ . The phosphorus isotope, ^{31}P , also exists in 100% abundance, and its nucleus has a spin of $I = \frac{1}{2}$, but its γ is much smaller. It also suffers from an efficient relaxation process called *chemical shift anisotropy* (CSA), which broadens its peaks (see Appendix A.2). The nuclei of the most abundant isotope of carbon, ^{12}C , is not NMR-active ($I = 0$), whereas that of the less abundant isotope (1.07%), ^{13}C , has $I = \frac{1}{2}$, and is active. Therefore, the ^{12}C carbons are often replaced with ^{13}C by isotopic

TABLE 1.17. Properties Determining Nuclear Spin^a

Mass Number	Atomic Number	Nuclear Spin (I)
Even	Even	0
Even	Odd	Integral (1, 2, 3...)
Odd	Even or odd	Half-integral ($\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \dots$)

^aWhen $I = \frac{1}{2}$, the nucleus has a spherical electric field, resulting in sharp peaks in the NMR spectrum. When $I > \frac{1}{2}$, the nucleus has an ellipsoidal electric field, resulting in quadrupolar relaxation and hence broad NMR lines. All nuclei with $I \neq 0$ are observable by NMR (i.e., are “NMR-active”).

TABLE 1.18. NMR Properties of Selected Nuclei^a

Nucleus	Nuclear Spin (<i>I</i>)	Number of Spin States ($m = 2I + 1$)	Gyromagnetic Ratio (γ) (T·s) ⁻¹	Natural Abundance (%)
¹ ₁ H	$\frac{1}{2}$	2	2.675×10^8	99.988
² ₁ H	1	3	4.107×10^7	0.012
¹² ₆ C	0	0	0	98.89
¹³ ₆ C	$\frac{1}{2}$	2	6.728×10^7	1.07
¹⁴ ₇ N	1	3	1.934×10^7	99.63
¹⁵ ₇ N	$\frac{1}{2}$	2	-2.712×10^7	0.37
¹⁶ ₈ O	0	0	0	99.96
¹⁷ ₈ O	$\frac{5}{2}$	6	-3.628×10^7	0.04
³¹ ₁₅ P	$\frac{1}{2}$	2	10.839×10^7	100.00

^aThe atomic number is given as a subscript and the atomic mass, as a superscript.

Source: Data taken from Ref. 204.

enrichment in biological samples. In the case of nitrogen, there are two NMR active isotopes, ¹⁴N and ¹⁵N, of which ¹⁴N is the most abundant, but its nucleus has a quantum spin number of $I = 1$. Nuclei with $I = 1$ produce spectra that are often difficult to interpret, and ¹⁴N has the added disadvantage of undergoing very fast quadrupolar relaxation, producing broad bands in the spectrum (see Appendix A.2.3). On the other hand, ¹⁵N, which is not as abundant (0.37%), has a spin of $\frac{1}{2}$ and has better relaxation properties. Once again, the concentration of ¹⁵N is usually enriched in biological samples to enhance its detection.

Bar magnets interact and align themselves with respect to each other. Similarly, when nuclei are placed in a strong, external magnetic field, the nuclear magnetic moments also align with that magnetic field. NMR spectroscopy is based on manipulating that alignment by irradiating the sample with radiofrequency pulses of appropriate energy. By pulsing the sample in different ways, a wide variety of experiments can be performed, each of which probes the molecule for different properties. The most common analyses made via NMR are chemical shift (indicative of the local chemical environment), resonance intensity (a measure of the relative numbers of a given type of atom), spin–spin coupling (a function of through-bond connectivity, and hence chemical structure), the nuclear Overhauser effect (a function of through-space proximity, and hence of molecular conformation), local and global dynamics as determined by relaxation properties (arising in part from the previous two), chemical exchange, and diffusion.

1.4.1. NMR Signals and Magnetization Relaxation

As noted above, the magnetic fields of nuclei align themselves in the presence of a strong magnetic field. The net magnetization of the nuclei can be expressed as a vector that points in the same direction as the magnetic field, which is defined as being along the z axis (Fig. 1.47). In a typical modern NMR experiment, the sample is briefly

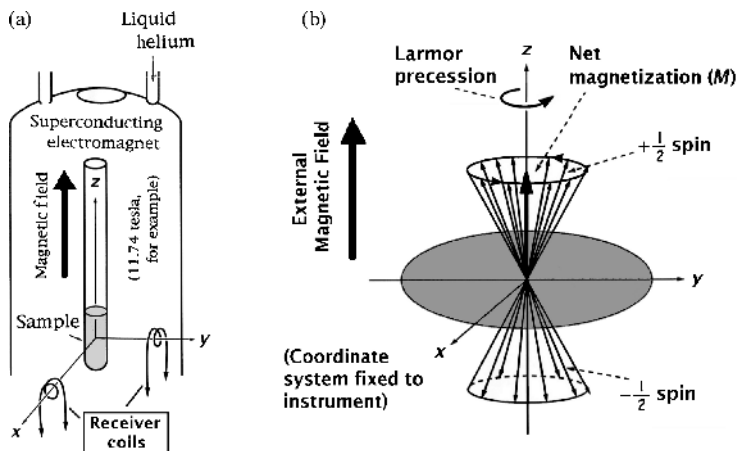


FIGURE 1.47. A comparison of the macroscopic and microscopic aspects of an NMR experiment with respect to the laboratory's physical coordinate system, in which the magnetic field defines the $+z$ axis. Diagram (a) depicts the placement of the sample and receiver coils within the instrument, and indicates that located along the x and y axes. Diagram (b) shows the orientations of the nuclear magnetic moments of the sample with respect to their spin states. There are more spins aligned with the external magnetic field (B_0) than opposed to it, hence the net magnetic vector is also aligned with it, and points in the $+z$ direction. The dipole moments of the nuclei rotate around the z axis with a periodicity proportional to the gyromagnetic ratio (γ). This is called the *Larmor precession frequency*. For a magnetic field strength of 11.74 T, a proton (^1H) will precess at a rate of 500 MHz.

irradiated with a radiofrequency (RF) pulse, which rotates the nuclei into the x,y plane (perpendicular to the z axis). The intensity of an NMR signal is determined by the amount of magnetization in the x,y plane as a function of time. When the RF pulse is turned off, the nuclei begin to return to the initial alignment via a variety of relaxation processes (Appendix A.2.3). These processes are depicted in Fig. 1.48, and fall into two general categories: longitudinal relaxation (*spin-lattice*) and transverse relaxation (*spin-spin*). Together, they give rise to all measurable NMR properties. For example, dipole-dipole relaxation, which is one of several longitudinal relaxation processes that involve the interaction of nuclei through space, gives rise to the nuclear Overhauser effect (NOE; see Section 1.4.4) as well as dipolar coupling (see Section 1.4.5). Similarly, transverse relaxation, which arises from spin-spin coupling (also called *scalar*, or *J coupling*), is the origin of the splitting patterns in one-dimensional (1D) spectra. It is this process that enables through-bond correlation experiments, which provide information on the connectivity of atoms within a molecule (e.g., COSY and HSQC; see Section 1.4.3).

These relaxation properties are often affected by experimental conditions, such as temperature and viscosity, which can alter the NMR spectrum, especially the peak locations and their linewidths. The former is illustrated in Fig. 1.49, which shows the phosphorus (^{31}P) NMR spectrum of $\text{d}(\text{CCTAGG})_2$ at 11°C , as well as a graph of how the positions of the peaks change with temperature [205,206].

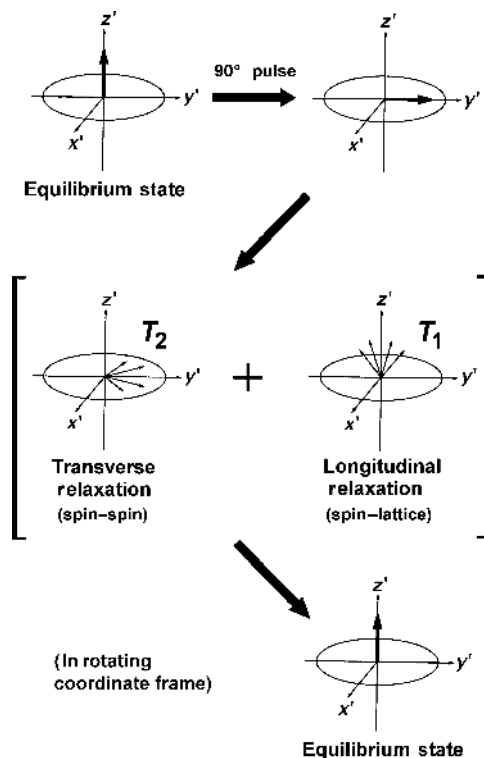


FIGURE 1.48. In their initial, equilibrium states, the nuclei are aligned such that their net magnetization is along the $+z$ -axis ($M_z \neq 0$), with no net magnetization in the x, y plane ($M_{x,y} = 0$). Then, the nuclei are depicted as being irradiated with a 90° pulse radiofrequency along the x axis, causing the net magnetization ($\mathbf{M}$) to rotate away from the z axis, and aligning it along the y axis ($M_y \neq 0, M_z = 0$). After the pulse is turned off, the magnetization relaxes back to its equilibrium alignment with the magnetic field via two relaxation processes. Longitudinal relaxation is the vertical realignment of $\mathbf{M}$ along the $+z$ axis ($M_z \neq 0$). Transverse relaxation is the restoration of zero net magnetization within the x, y plane ($M_{x,y} = 0$). The latter is achieved by the radial dispersion of the individual nuclei within the x, y plane, resulting in a net cancellation of the magnetization. These processes produce all measurable NMR phenomena and result in the free induction decay (FID) curve. The coordinate system ($x' y' z'$) is rotating with respect to the fixed coordinate system ($x y z$) around the z axis. See text and Appendix A.2, Section A.2.3, for more information.

In addition, these relaxation properties can provide useful information about the dynamical motion, or conformational fluctuation, within a molecule or its complex. Such motion is usually altered when a drug or other ligand binds to DNA, and can sometimes be correlated with the preferential binding of the drug at certain base sequences. For example, an early study of the flexibility of calf thymus DNA (CT-DNA) measured the transverse relaxation time of its ^{31}P NMR spectrum [207].

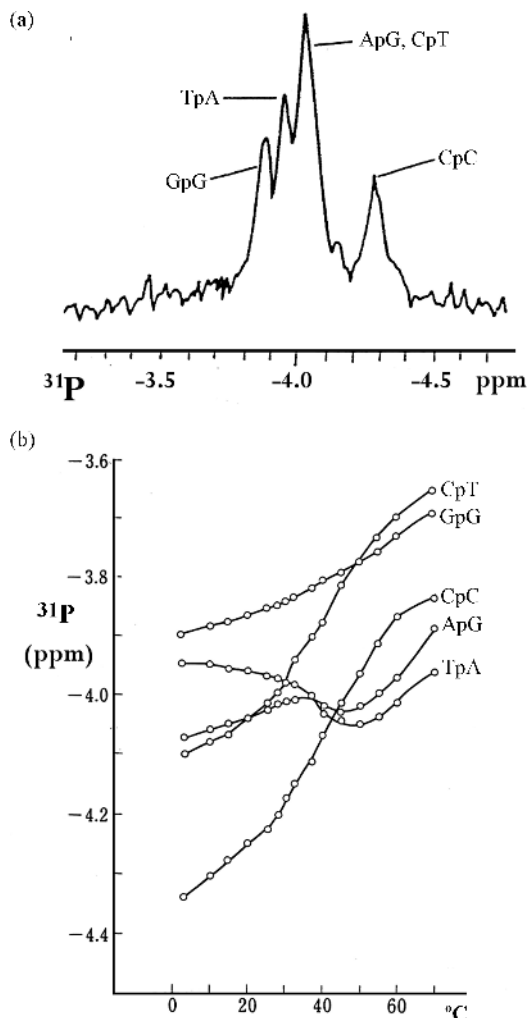


FIGURE 1.49. (a) The phosphorus NMR spectrum of $\text{d}(\text{CCTAGG})_2$; (b) ^{31}P chemical shifts of $\text{d}(\text{CCTAGG})_2$ as a function of temperature and their assignments [205].

More information about the mechanisms and measurement of relaxation properties can be found in Appendix A.2.3.

1.4.2. Proton NMR Spectra of Bases, Nucleosides, and Nucleotides

1.4.2.1. Chemical Shifts Figure 1.50 is an example of a proton (^1H) NMR spectrum of a nucleoside. This is a spectrum of cytidine that was obtained in a D_2O solution, using a spectrometer with a magnetic field strength of 5.87 T (teslas) [208]. Cytidine has 12 hydrogen atoms, each of which is observed in the NMR spectrum as a

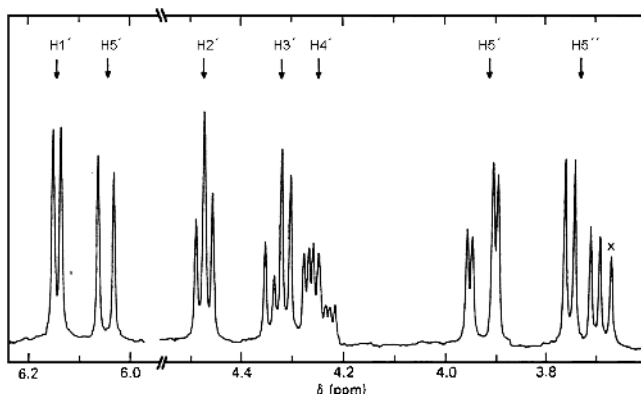


FIGURE 1.50. The 250-MHz ^1H NMR spectrum of cytidine in $^2\text{H}_2\text{O}$ [208].

peak occurring at a different resonance frequency, or “chemical shift” (δ), expressed in units of ppm (see Appendix A.2.2). These different frequencies arise from the fact that each proton (or ^1H hydrogen atom) is in a slightly different electronic environment, and hence “feels” the external magnetic field slightly differently. These resonances can be summarized as follows:

Base protons:	H5: $\delta = 6.049$ ppm;	H6: $\delta = 7.752$ ppm (not shown);
Sugar Protons:	H1': $\delta = 6.144$ ppm;	H2': $\delta = 4.478$ ppm;
	H3': $\delta = 4.310$ ppm;	H4': $\delta = 4.254$ ppm;
	H5': $\delta = 3.930$ ppm;	H5'': $\delta = 3.731$ ppm.

Nuclei in similar environments always have comparable resonance frequencies and generally fall within a relatively small range of ppm values [198,209,210]. These vary systematically as a function of nucleotide sequence and global structure, arising from differences in the local chemical environments. For example, the H8 of deoxyguanosine in the second G (or G_2) of the sequence $\text{d}(\text{CGCAGAATTCGCG})_2$ has $\delta = 7.95$ ppm, whereas the H8 of the 12th base (or G_{12}) has $\delta = 7.76$ ppm. Thus, the effect of DNA sequence and base-pairing influences the observed resonance frequencies, but the different DNA residues can still be distinguished. Table 1.19 lists the chemical shift values of hydrogen, carbon, and nitrogen of some nucleoside residues in B-form DNA.

1.4.2.2. Spin–Spin Coupling Constants The ^1H NMR spectrum of cytidine (Fig. 1.50) also demonstrates another important feature of NMR spectra—the resonance lines for most of the hydrogen atoms do not appear as single peaks, but rather are split into multiplets. The separation between these peaks gives the *coupling constant* J , and is measured in Hertz (Hz). The physical mechanism for this effect is the transfer of magnetization via the spin pairing of each atomic nucleus to the spins of the electrons in chemical bonds between them. It is thus called an *electron–coupled, spin–spin*

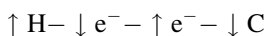
TABLE 1.19. Normal Ranges of Proton (^1H) and “X”-Atom [^{13}C] or Nitrogen (^{15}N)] Resonance Frequencies in a Standard DNA Duplex

Atom Position/Group	Resonances (ppm)	
	^1H	X
Deoxyribose		^{13}C
1'	5.2–6.7	84–91
2' (H', H'')	0.9–3.9	37–44
3'	4.1–5.6	72–82
4'	3.8–5.0	84–89
5' (H', H'')	3.3–4.6	62–72
Nucleoside ^b		^{13}C
2 (A)	7.3–8.4	153–155
5 (C)	4.5–6.5	92–99
6 (C, T)	6.5–8.1	139–143
7 (T)	0.8–1.6	13.5–15
8 (A, G)	7.4–8.5	137–142
		^{15}N
–NH ₂ (A, C, G)	6.2–9.0	(A) 79–81 (G) 74–76 (C) 96–98
>NH (ring NH of G, and T)	10–15 m	(G) ~147 (T) 158–160

^aThese numbers were compiled from multiple sources [198,209]. The ranges listed here are approximate, as many outlying examples exist. In particular, the resonances arising from cytidine may have a much wider range than those given here, as they may exist in both neutral and acid forms.

^bNotice that the chemical shifts of aromatic atoms in the nucleosides are much higher than those in deoxyribose moiety.

interaction, which can be depicted for a C–H bond as



where the arrows refer to the orientations of the quantum spin states. Thus, each nucleus is sensitive to the possible orientations of the spin of its neighbor. This method of transferring magnetization is an important transverse relaxation process in NMR called T_2 (see Fig. 1.48, left, Section 1.4.3, and Appendix A.2, Section A.2.3). The strength of the coupling between two nuclei i and j is given by the magnitude of the interaction energy

$$E = \hbar J(\mathbf{I}_i \cdot \mathbf{I}_j), \quad (1.6)$$

where $\hbar$ is Planck's constant, $\mathbf{I}_k$ is the spin quantum number of nucleus k , and J is the *coupling constant*. This coupling is also called *J coupling* or *scalar coupling* because of the form of this mathematical equation.

Under normal circumstances, this spin–spin coupling extends out to about three bonds (e.g., H–C–C–H), although coupling through more bonds is possible. This *through-bond* coupling provides information on an atom’s chemical environment, and it can be exploited to determine unambiguously the connectivity of a chain of atoms by using one of the spectroscopic, 2D correlation experiments, such as COSY (Section 1.4.3).

When a nucleus is coupled to a neighbor with a nuclear spin of $I = \frac{1}{2}$, its resonance will be split into a doublet; if the neighboring atom has a spin of $I = 1$, then it will be split into a triplet. Thus, the general rule is that the number of lines in a multiplet is $(2nI + 1)$, where n is the number of identical atoms in a neighboring chemical group and I is the spin of these nuclei. The intensity of each line in the multiplet is determined by the number of possible combinations of individual spins that can produce a given value of total spin, and follows a binomial distribution. That is, when a nuclei is coupled to a group of n identical protons, its resonance is split into a series of peaks whose number and intensity ratios are given in the n th row of Pascal’s triangle (for couplings to nuclei with $I = \frac{1}{2}$):

n	Resonance Pattern			Splitting Pattern
0	1			Singlet
1	1	1		Doublet
2	1	2	1	Triplet
3	1	3	3	Quartet
	(etc.)			

The situation can become quite complex if a nucleus is coupled to several non-equivalent nuclei, or if there is longer-range coupling, such as through a double bond.

Returning to the cytidine spectrum as an example (Fig. 1.50), the H1’ has an $I = \frac{1}{2}$ neighbor, H2’, and hence has its resonance line split into a doublet with $J = 3.2$ Hz. The H2’, however, has a deceptively simple triplet pattern. In fact, this is not a triplet at all, but a doublet of doublets in which the two inner lines overlap. This happens because H2’ is coupled with both H1’ and H3’, and the coupling constant is nearly the same with each, $J = 3.7$ Hz. This results in the same intensity pattern as the $n = 2$ case, but really results from two $n = 1$ groups.

It is possible to predict these coupling constants by using the Karplus equation [211], which is an experimentally parameterized relationship between the magnitude of the coupling constant, J and the dihedral torsion angle ϕ about the central bond in a chain of four atoms. This equation has several forms depending on the number of dimensions in the spectrum being analyzed. For a 1D NMR spectrum, it is

$$J(\phi) = A \cos^2(\phi - 60) - B \cos(\phi - 60) + C \quad (1.7)$$

where A , B , and C are empirical parameters, which depend on the types of atoms involved in the coupling. This relationship between coupling constant and dihedral torsion angle (ϕ) is graphed in Fig. 1.51 for HCCH coupling. It clearly reveals that these coupling constants are largest when ϕ is either 0° or 180° , and smallest at 90° .

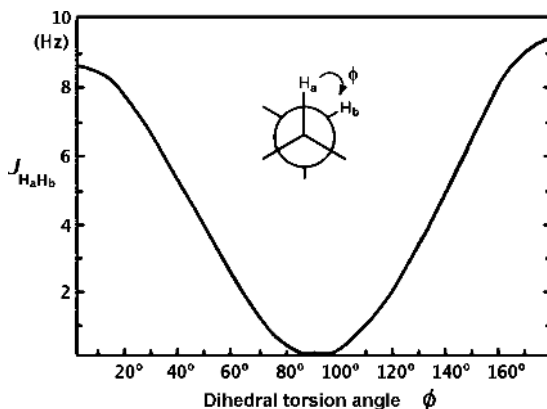


FIGURE 1.51. The Karplus curve; showing dependence of the coupling constant J (in Hz) on the dihedral angle ϕ , which is defined by the planes of $H-C_i-C_j$ and C_i-C_j-H bonds in an $H-C_i-C_j-H$ fragment.

This equation is very important when determining 3D structures of biological macromolecules from NMR data.

This mathematical function applies only to weakly coupled systems, where J is much smaller than the chemical shift difference between the coupled nuclei (and all frequencies are expressed in units of Hertz); that is, it applies when $10|J| < \Delta\delta$, where $\Delta\delta$ is the difference in the chemical shifts, for example $(\delta(H_i) - \delta(H_j))$ in Hertz. This defines a “first-order” spectrum. If J has at least the same order of magnitude as $\Delta\delta$, then more complex spectral patterns are observed, as in the 4.2–4.4 ppm region of the cytidine spectrum (Fig. 1.50). As the external magnetic field strength B_0 increases, the chemical shift difference ($\Delta\delta$, in Hz) between the coupled peaks is also increased. Since almost all modern spectrometers operate above 400 MHz for the ^1H , $\Delta\delta$ is greater than $10|J|$. This means that all protons are weakly coupled, and their spectra can now be considered to be first order.

1.4.3. Correlation Analysis, or “Spin–Spin Coupling”

When determining the three dimensional structure of a molecule, it is necessary to first assign the NMR resonances to specific atoms within the molecule. An atom’s chemical shift value (ppm) often provides a general classification of its chemical environment, but it gives no information regarding the atom’s connectivity (bonding) to other specific atoms. As noted in the previous section, the nuclear magnetic dipole of one atom can interact with a chemically attached neighboring atom through their shared electrons in the chemical bond. This is a T_2 relaxation process (see Appendix A.2.3) and there are many NMR experiments that take advantage of this to determine chemical connectivity.

For example, chemical connectivity may be determined by a one-dimensional experiment in which a sample is irradiated with a radiofrequency equal to the

resonance of a particular atomic nucleus. This “decouples” any other resonances that are spin-coupled to that irradiated nucleus, resulting in the collapse of the coupling patterns. Hence, irradiating the H2' of cytidine at 4.48 ppm will cause the H1' doublet at 6.14 ppm to collapse into a singlet located at its midpoint (compare with Fig. 1.50). This radiofrequency causes the irradiated nuclei to rapidly transit between their possible spin orientations, rendering the different spin states indistinguishable [193].

By using a variety of such pulse sequences, it is possible to take advantage of these T_2 relaxation processes in order to transmit magnetization information along a chain of atoms of two bonds, or more. This can unequivocally establish a series of atoms that are chemically connected to each other. Common two-dimensional experiments that measure this are COSY, TOCSY, HMBC, HSQC, HMQC, and HETCOR (or heteronuclear COSY). The first two of these are homonuclear experiments that establish through-bond connectivities between protons. The remaining four experiments establish through-bond connectivities between protons and heteroatoms (e.g., ^{13}C , ^{15}N , and ^{31}P). In addition to these experiments, many others exist that have been designed specifically for certain chains of atoms, especially those found in proteins. We will not consider all of these different experiments, but instead only discuss the uses of general types.

1.4.3.1. Homonuclear Correlations and Coupling Analysis The homonuclear COSY experiment establishes which protons (^1H) are on the same or adjacent carbon atoms, but it provides no direct information about the intervening carbon atoms. In nucleic acids, the ^1H - ^1H coupling patterns within the sugar rings are particularly informative, as they can allow one to unequivocally distinguish most of their resonances. The TOCSY experiment provides similar information as the COSY, but it relays magnetization through an entire network of spin-coupled nuclei. So, the COSY shows interatom connectivity between adjacent protons on a pairwise basis, whereas the TOCSY reveals all protons in a group. Together, these experiments can be used to assign the signals of all the sugar protons in a nucleotide.

The phase-sensitive variant of the COSY (called DQF-COSY) contains additional information about the J coupling between protons. It is often used to determine the sugar pucker conformations of the nucleotides. This is possible because each proton in a sugar ring is coupled to its adjacent neighbor, and the coupling constant between protons is an indication of the dihedral torsion angle (ϕ), as given by the Karplus equation (Section 1.4.2). Thus, by carefully determining the various coupling constants, one is able to determine the pseudorotation angle P (Section 1.1.4) [212–216]. These angles can, in turn, be correlated to the backbone torsion angles. Although several programs are available for this type of detailed analysis, it remains fairly tedious. It is also often possible to make a more qualitative classification of the pseudorotation angle by simply comparing the experimental coupling patterns of the H1'-H2' and H1'-H2'' COSY cross-peaks to those found in the literature. If more than one conformation exists for a molecule, or if a sugar undergoes rapid dynamical motion, such as rapid flipping of its pseudorotational angle (“repuckering”), then these coupling patterns can become diffuse, or indistinct, or their different phases may result

in cancellation of the peak intensities. There are other techniques for obtaining this information that do not rely on COSY information. For example, in RNA studies, it is often sufficient to ascertain the strength of a H1'-H2' cross-peak in a 2D-NOESY experiment (Section 1.4.4). Reliable information about sugar conformations can be also obtained from *residual dipolar coupling* analysis (Section 1.4.5).

1.4.3.2. Heteronuclear Correlations and Coupling Analysis Heteronuclear correlation NMR experiments are quite important when determining the sequential base connectivity. The prototype of these is the heteronuclear COSY experiment (also called "HETCOR" or [$^1\text{H}, \text{X}$]-COSY), which can determine which proton is coupled to a nearby heteroatom X, where $\text{X} = ^{13}\text{C}$, ^{15}N , or ^{31}P . An example of a heteronuclear C,H-COSY experiment of the self-complementary DNA dedecamer, $\text{d}(\text{CGCGAATTCGCG})_2$ is shown in Fig. 1.52. In the seventh residue of this duplex, thymidine-7 (T_7), the carbon atoms of the deoxyribose sugar were uniformly, isotopically enriched with ^{13}C [217]. This enhanced the carbon signals from this residue. These maps show cross-peaks ("spots") for each C-H bond, giving the chemical shifts of each atom in the bonding pair, as well as allowing the measurement of their spin couplings. More sensitive experiments, such as HSQC or HMQC, give the same kind of information and have been used in more recent publications.

If a sample has been properly enriched, it is also possible to perform heteronuclear experiments of the type [X,Y]-COSY, where X and Y can be any NMR-active isotope (e.g., ^{13}C , ^{15}N , or ^{31}P). For example, all of the atoms within the aromatic rings of a nucleotide might be assigned by using an [^{13}C , ^{15}N]-COSY, or [^{13}C , ^{15}N]-HSQC experiment.

1.4.3.3. Detection of Hydrogen Bonding in Base Pairs Scalar coupling NMR experiments can directly detect the imino hydrogen bonds in DNA and RNA base pairs. These imino hydrogen bonds occur between two nitrogens, one of which is a hydrogen donor and the other is a hydrogen acceptor: $\text{N}-\text{H} \cdots \text{N}$. For this experiment, the aromatic bases of the DNA are enriched with ^{15}N to enhance their detection, and the nitrogens are assigned by means of a heteronuclear correlation experiment, such as an HSQC (see discussion above). Then, by using another scalar coupling experiment (such as HNN-COSY), one can measure the couplings between the nitrogens *through* the hydrogen bond of the base pair, $^{15}\text{N}-^1\text{H} \cdots ^{15}\text{N}$ [218,219].

These types of scalar experiments provided the first direct proof of such hydrogen bonds in nucleic acids. Prior to this, the presence of hydrogen bonds could be inferred only from the geometries of donor and acceptor groups found in 3D models of nucleic acid structures, or from NOE measurements (Section 1.4.4.2). These scalar couplings are also particularly important as they allow the direct observation of unique base-pairing patterns, such as Hoogsteen base pairs [220, 221]. The N-N couplings appear to be very sensitive to the atomic level details of the hydrogen bonds, as there is evidence of a correlation between the size of the scalar coupling and the strength of the hydrogen bond [222].

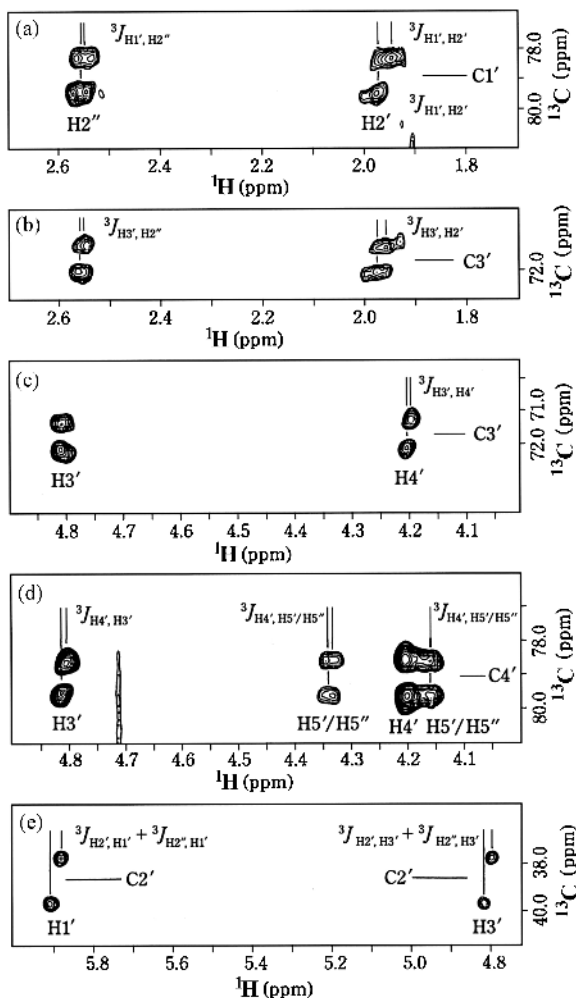


FIGURE 1.52. Two-dimensional maps of spin–spin couplings between ^{13}C and ^1H in d(CGCGAATTCGCG) $_2$ from a C,H–COSY. The seventh base, T $_7$, is a thymidine in which the deoxyribose carbons (C1', C2', C3', C4', C5') were uniformly enriched with the isotope ^{13}C [217]. The chemical shifts were determined to be 81.0 (C1'), 38.6 (C2'), 71.6 (C3'), and 79.1 (C4') ppm, and their couplings measured: (a) C1'–H2'/H2'' couplings; (b) C3'–H2'/H2'' couplings; (c) C3'–H3'/H4' couplings; (d) C4'–H5'/H5''/H3' couplings; (e) C2'–H1'/H3' couplings.

1.4.3.4. Phosphorus–Proton Couplings and Backbone Dihedral Angles

For nucleic acids, a particularly useful heteronuclear experiment is the [^1H , ^{31}P]-COSY (HETCOR), as it can establish the sequential connectivities between sugar moieties. It does this by identifying the coupling between the H3' protons of base n , through to the ^{31}P in the phosphodiester linkage, and then from the ^{31}P to the H4' or

H5'/H5'' of the next base, ($n + 1$). The H5'/H5'' peaks are usually weaker and more difficult to resolve; hence H4' is more commonly utilized.

As an example of the utility of ^{31}P NMR in the study of short DNA oligomers, the 2D [^1H , ^{31}P]-COSY spectrum of d(CCTAGG) $_2$ is shown in Fig. 1.53 [206]. By comparing this spectrum with the resonance assignments for H5' and H4' (determined from ^1H experiments), the assignments of the ^{31}P signals were established. This approach is often used to confirm, or determine, the sequential assignment of the DNA resonances. The different chemical shift values of these phosphorus atoms reflect their different environments within the polynucleotide chain. It was noted that the ^{31}P chemical shift has a nearly linear correlation to the local twist angle t_g of the polynucleotide chain [207]. Here, t_g is the change in the orientation of the C1' – C1' vectors in two successive base-pairs when viewed down the helical axis (the standard B structure is assumed to have a value of $t_g = 35.9^\circ$; see Table 1.1 and Section 1.1.8).

These P–(O–C)–H couplings are also dependent on the dihedral torsion angle and can be used to determine the backbone conformation of an oligonucleotide by an equation similar to the Karplus equation [223–231]. For example, $J_{\text{PH}} = 28$ Hz corresponds to the *trans* conformation, while $J_{\text{PH}} = 1.5$ Hz corresponds to the *gauche* conformation. Unfortunately, this is not a very sensitive technique, making such analyses difficult. In addition, the chemical shift dispersion of ^{31}P is usually small (only a couple of ppm units), which limits the utility of this approach to smaller oligonucleotides (less than ~ 20 residues).

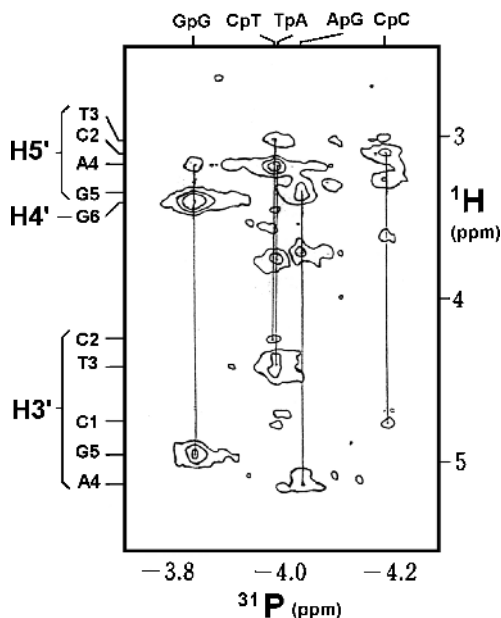


FIGURE 1.53. A contour plot of a ^{31}P - ^1H shift correlated two-dimensional NMR map [^{31}P , ^1H]-COSY of d(CCTAGG) $_2$ (1400 OD/mL) at 31°C [206].

In some cases one can obtain reasonable ranges of the backbone torsion angles based on the combined use of linewidth analysis and the sums of the J couplings for different sugar protons [232]. The sum of the $H4'$ J couplings

$$\Sigma J_{H4'} = (J_{H4'-H3'} + J_{H4'-H5'} + J_{H4'-H5''}) \quad (1.8)$$

is dependent on the values of the sugar pseudorotation angle P and the backbone angle γ (see Figs. 1.2 and 1.15). $J_{H4'-P}$ depends on both γ and β , but its dependence on γ is ambiguous. Similarly, the sum of the $H3'$ J couplings

$$\Sigma J_{H3'} = (J_{H3'-H2'} + J_{H3'-H2''} + J_{H3'-H4'}) \quad (1.9)$$

is dependent on the backbone angle ϵ . Some of the other backbone angles are better estimated from NOE distance information (Section 1.4.4).

In the discussion above, only the heteronuclear COSY experiment (HETCOR) was mentioned, but there are a number of more sensitive pulse sequences that give similar information, such as the HSQC and HMQC experiments. These experiments can provide the same sequential connectivity information as the HETCOR. However, even with these experiments it can be difficult to obtain such information because of spectral overlap and the fast relaxation properties of ^{31}P nuclei. As it is not always possible to obtain all this information, spectroscopists sometimes use general estimates of the backbone conformation, which serve to keep the molecular models physically reasonable, without unduly biasing the final result.

A number of synthetic modifications have been suggested as a means for learning about the nucleic acid backbone connectivity and conformation, without relying on ^{31}P . For example, oligonucleotides can be synthesized with the phosphodiester linkage replaced with a phosphoramidate linkage. Here, one of the phosphoester oxygen atoms is replaced by ^{15}NH . Although this substitution can be done with the oxygen atom on either the 3' or 5' side of the P atom, only the -C3'-NH-P5'-O-C5' variant forms a stable duplex (compare with Fig. 1.2). This is the N3' $\rightarrow$ P5' derivative, and it prefers to stabilize the A form over the B form in both DNA and RNA [233]. Since RNA prefers the A form, this modification is generally more effective in RNA studies.

1.4.4. Nuclear Overhauser Effect

The *nuclear Overhauser effect* (NOE) enables the measurement of proton–proton interatomic distances through space. This makes it extremely important when studying biomolecules, and is especially critical in the determination of nucleic acid structures. This effect occurs when there is a transfer of magnetization between two nuclei through space via the dipole–dipole relaxation process (see Appendix A.2.3), and it is observed as a change in the intensity of an NMR signal. An NMR signal can be enhanced by a positive NOE, or decreased by a negative one. Three primary factors affect the sign and intensity of an NOE: (1) the gyromagnetic ratios of the interacting nuclei, (2) the interatomic distance between the nuclei, and (3) the mobility of the

molecule (or of a substructure within the molecule). This can be expressed by the relationship [192,193]:

$$I_{\text{NOE}} \propto \frac{\gamma_j}{\gamma_i} \frac{1}{r_{ij}^6} \cdot f(\tau_c) \quad (1.10)$$

where i is the observed nucleus, j is the interacting nucleus, r_{ij} is the distance between the two nuclei, γ is the gyromagnetic ratio of a nucleus, and $f(\tau_c)$ is a complex function of the correlation time of the molecule τ_c . The correlation time is a measure of the tumbling rate of the molecule in solution (Appendix A.2, Section A.2.3), and is, itself, a function of many factors, including the size and shape of the molecule, especially its molecular weight. The correlation time is therefore a description of molecular mobility.

Since the NOE is dependent on (γ_j/γ_i) , it is clear that the relative values of γ_i and γ_j affect the intensity of the NOE. It is also clear that the signs of these gyromagnetic ratios can determine the sign of the NOE. When both nuclei have positive values of γ (e.g., ^1H or ^{13}C), then the NOE is also positive and the NMR signal is enhanced. When one of the nuclei has a negative value of γ (e.g., ^{15}N), then the NOE is negative, and the NMR signal is decreased.

The correlation time τ_c also affects the sign and magnitude of NOEs. This influence can be seen by plotting NOE intensity as a function of molecular weight (a primary determinant of τ_c ; see discussion above). Small molecules, such as many drugs, tumble rapidly in solution, resulting in small τ_c values and hence small, *positive*, NOEs. Large macromolecules, such as DNA, tumble slowly. These have larger τ_c values and large, *negative* NOEs. In the middle are “medium-sized” molecules, with molecular weights corresponding to intermediate values of τ_c . Such molecular weights fall near the point where the NOEs change from positive to negative. As a result, medium-sized molecules have NOEs that are small or nearly zero. This can be a problem when trying to determine 3D structures of such molecules. Consequently, experiments were developed that measure the *rotating-frame nuclear Overhauser effect* (ROE), instead of the NOE. ROEs provide the same kind of information as NOEs, but they have the same sign regardless of the molecular weight. Whereas the NOESY experiment measures NOEs, the ROESY measures ROEs. The ROESY experiment can also give information about molecules that are in exchange between different conformations in the sample. The description of the ROESY experiment is more complicated than the NOESY, and it will therefore not be discussed in detail here.

Perhaps the most important aspect of Eq. (1.10) is that the magnitude of NOE (and ROE) enhancement is inversely proportional to the distance between the atoms (r_{ij}), raised to the sixth power: $I_{\text{NOE}} \propto (1/r_{ij})^6$. This allows the calculation of interatomic distances based on NOE intensities. The inverse, sixth-power relationship means that NOE intensity diminishes quickly as the distance between the atoms increases and indicates that less magnetization is transferred at greater separations. Thus, NOEs are quite strong when atoms are $<2 \text{ \AA}$ apart, but they become very small when the atoms are separated by more than $\sim 5 \text{ \AA}$. This means that NOEs can be used only to determine short distances.

As with the scalar coupling experiments, NOEs may be measured by either a series of 1D experiments, or by a single 2D experiment. In the standard 1D difference experiment, the magnitude of an NOE is determined one resonance at a time. A steady-state protocol is used in which the sample is selectively irradiated with a radio-frequency that is equal to the energy of the resonance signal of interest. This saturates the irradiated spin—that is, it equalizes the populations in all the possible spin states. This, in turn, perturbs the spin populations of any neighboring nuclei that are dipole-coupled to the irradiated nucleus. A control spectrum is then acquired in which the irradiation is set to a frequency far away from any resonances. Any intensity differences between these two spectra (either positive or negative) may be considered to be the result of an NOE between the two nuclei. Using this information, one can identify the resonances of nuclei that are within 5 Å of each other. As the approach described here uses a steady-state protocol, it is able to give only relative internuclear distances.

The 2D NOESY experiment uses a transient protocol that starts by moving the net magnetization of the nuclei into the $-z$ axis. This is a simultaneous inversion of all the spin-state populations, which allows all possible pairs of nuclei to interact. Then, the buildup of the NOE signals is measured as the nuclei return to the $+z$ axis via *longitudinal relaxation* (see Fig. 1.48 and Appendix A.2.3). This procedure provides accurate internuclear distances [193]. Its spectral resolution in either of its dimensions is not as great as the 1D variant, but this is usually more than compensated by the dispersion into a second direction. The 2D NOESY experiment is generally more visually intuitive to interpret, as it can be regarded as a “contact map” in which only protons that are close to each other show a cross-peak (a spot). Although NOESY data are not acquired as a series of 1D NOE experiments, the analysis can be performed as if the NOEs were measured one resonance at a time.

1.4.4.1. The NOESY-Walk The NOESY experiment provides the basis for the “NOESY-walk” analysis of nucleic acids. The NOESY-walk takes advantage of the fact that in DNA, certain pairs of inter-residue protons are almost always within 5 Å of each other. The most important “walk” is between the aromatic protons of the bases and the H1' of the sugar moieties. The intra-residue $H_{\text{aro}}-\text{H1}'$ distance is typically $\sim 3.6\text{--}3.9$ Å (for the *anti* glycosidic dihedral torsion angle), resulting in a medium-intensity cross-peak. The inter-residue NOESY cross-peaks between the aromatic protons on base n , and the H1' of base $(n - 1)$, are usually weaker, being separated by $\sim 4\text{--}4.9$ Å (see discussion below and Fig. 1.54c.). Together these peaks normally enable one to draw a connectivity chain from one end of the DNA strand through to the other end. The other sugar protons are also used for NOESY-walk analysis, as are the distinctive thymine methyl protons. Each of these “walks” contributes additional structural details, and provide informational redundancy that can help resolve ambiguous overlaps of NMR resonances in other parts of the spectrum. Often, the most revealing information comes from an unexpected “break” in a walk sequence, or when additional cross-peaks are observed between two adjacent aromatic bases or sugar residues. All of this information is compiled into a large table of short-range (≤ 5 Å) interproton distances, which can then be used with computational structure building programs to develop a high-resolution model.

As an example of a simple NOESY-walk, the $H1'$ -(aromatic base) portion of such a spectrum is shown in Fig. 1.54 [234] for the sequence d(TGGGGT). This DNA hexamer most likely forms a quadruplex with four equivalent, and indistinguishable, strands. In the one-dimensional ^1H NMR spectrum (Fig. 1.54a), four downfield H8 resonance lines are observed, corresponding to the four guanosines in each strand. There are also two upfield H6 resonance lines from the two thymines. The two-dimensional spectrum (Fig. 1.54b) is displayed with the aromatic region plotted

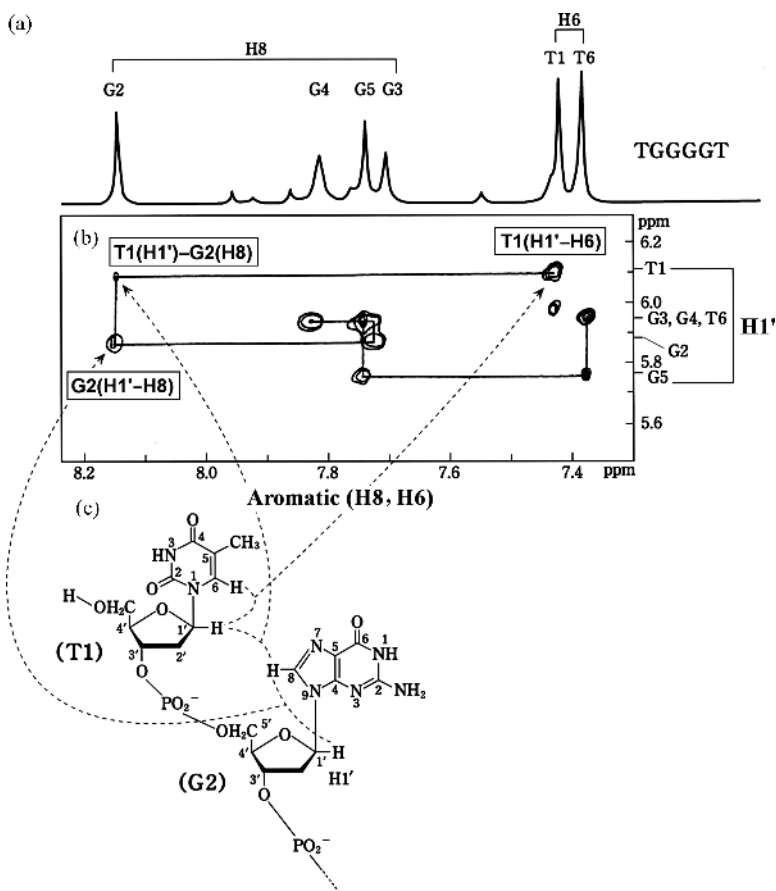


FIGURE 1.54. The one- and two-dimensional proton NMR spectra of the hexamer oligonucleotide, d(TGGGGT) [234]. This sequence mostly likely forms the linear DNA quadruplex, d(TGGGGT)₄, consisting of four strands in identical conformations (Section 1.1.12): (a) one-dimensional NMR spectrum in the resonance region of aromatic protons (8.2–7.4 ppm); (b) two-dimensional NOESY spectrum showing a map of the nuclear Overhauser effects between the aromatic protons (G-H8 and T-H6) of the bases and the $H1'$ of the deoxyribose rings; (c) a part of the molecular structure. The protons giving rise to each of the NOESY cross-peaks are indicated by dotted lines, and the associated peaks are shown with arrows.

across the horizontal axis and the H1' region plotted along the vertical axis. It is clear from Fig. 1.54(b, c) that the *intra*-residue distances between the H1' of residue T₁ and its own aromatic H6 are very short, as the intensity of the cross-peak is relatively strong. Likewise, the H1' of *intra*-residue G₂ is near its own aromatic base, H8. What makes this experiment particularly useful, however, is the fact that one also observes an *inter*-residue cross-peak between the H1' of T₁ and the H8 of G₂. This is not a strong peak, indicating that these two protons are separated by $\sim 4\text{--}5\text{ \AA}$. Thus, by using alternating *intra*- and *inter*-residue H1'-base cross-peaks, one can “walk” from the first residue of the strand (at the 5' end), through to the last base (3' end). This “walk” can be confirmed by the other possible “walks” mentioned earlier. Not only do these NOE cross-peaks allow the assignment of most of the aromatic and sugar resonances, but also, their intensities can be used to determine their approximate interproton distances, which are a measure of the proximity of the bases themselves.

An important indicator of an *intercalating* drug is that it often causes a conspicuous break in the NOESY walk, as the drug causes the DNA base step to expand at the site of the intercalation. If the drug molecule is tightly bound to the DNA, there also will be NOE cross-peaks between the drug and the DNA, especially between the aromatic or exchangeable protons of the base and the drug. Depending on how tight is the binding, it may even be possible to observe two sets of peaks, arising from both bound and unbound molecules. These intermolecular NOEs arise from the saturation transfer of magnetization between the molecules (see discussion below).

If two protons are not near each other, they do not give rise to a cross-peak in a NOESY spectrum. This “negative” information can be important, and it is often used to disprove certain molecular conformations. However, the absence of a peak in a NOESY spectrum is not always conclusive proof, as several factors can give false negatives. If a portion of the molecule is in dynamical motion (with respect to the NMR timescale), or if multiple conformations exist, the resulting peak may be too broad and diffuse to be observed. Likewise, if a drug is in rapid equilibrium (again, with respect to the NMR timescale), then it will also not produce a peak in a standard NOESY spectrum, although other approaches can be used (see text below).

1.4.4.2. Base-Pair Determination and Exchangeable Protons It is often necessary to determine the hydrogen bonding pattern in a DNA or a DNA–drug complex. In most NMR studies of nucleic acids, the solvent chosen is D₂O. This has the advantage of being an invisible solvent in proton NMR experiments, allowing the dense region of sugar protons to be observed near the water peak. However, hydrogen bonds can not be observed in this way because the hydrogen atoms are labile—they are in rapid exchange with the solvent and are quickly replaced with the ²H (D) atoms that have no signals in the proton spectrum. In order to observe these hydrogen bonds, the sample is dissolved in a solution of 10% D₂O and 90% H₂O. This mixed solvent provides a strong proton signal (¹H), enabling the observation of the hydrogen bonds. All of the standard 2D NMR experiments can then be performed, especially the 2D NOE. This allows one to observe the proximity of an N–H to other protons in nearby bases and sugar groups, and provides important interbase structural information,

especially for nonstandard base pairing. Imino protons are usually sharper, more easy to analyze and are more likely to give the desired structural information. Solvent suppression is required for these experiments, as the solvent HOD peak would otherwise overwhelm the digitized signal. The standard presaturation approach for suppression of the solvent signal cannot be used because the rapid exchange between the exchangeable protons and the water peak results in the suppression of both peaks [235].

Only those exchangeable protons that are protected from exchange with the solvent are observed. Protected protons are usually involved in hydrogen bonding, although some exchangeable protons are protected by virtue of being buried within a hydrophobic portion of a molecule, slowing their exchange rate with the solvent. Hydrogen-bonded protons are typically more downfield (have higher ppm values) than expected, as they are deshielded by the electronegative atoms (e.g., N and O) between which they are bound. These heavy atoms enhance the positive charge on the proton by pulling some of its electron density away: $N^{\delta-}-H^{\delta+}\cdots O^{\delta-}$, where δ denotes a partial electron charge.

Several examples of base-pairing schemes are depicted in Fig. 1.55, along with potentially observable interproton distances (represented as arrows in the figure) that are distinctive for that particular base pair. As shown in Fig. 1.55b, a NOESY cross-peak is expected between the N3H of thymine (or uracil) and the C2H of adenine. For the G-C base pair (Fig. 1.55a), on the other hand, an NOE signal is expected between the N1H of guanine and the NH₂ of cytosine, although in practice it is difficult to observe. It is also possible to observe sequential NOE contacts between adjacent DNA base pairs [198,236]. This can be used to confirm sequential assignments, which is especially important in nonstandard DNA structures (see, e.g., Ref. 237). Also, it is quite common to observe intermolecular NOE cross-peaks between these labile DNA-base protons and a drug molecule. These may arise with both *intercalating* drugs and *groove-binding* drugs. In the latter case, the drugs usually form hydrogen bonds with the exposed hydrogen bond acceptors and donors of the base pairs (for examples, see Chapter 3).

Besides the Watson-Crick A-T and G-C base pairs (Fig. 1.10), other special base pairings, called *Hoogsteen* and *reversed Hoogsteen base pairs*, [see the A-U variants in Fig. 1.55, parts (c) and (d), respectively] have been found [238]. The NOE patterns of these base pairs differ from those of the Watson-Crick pairings, as the N3H of uracil has a NOE signal to the C8H of adenine, instead of C2H. A standard Hoogsteen pairing is expected to have a weak (~ 4.4 Å) NOE between the adenine amino protons (N6H) and the uracil C5H. A reversed Hoogsteen pairing might also have weak NOE contacts between the adenine amino protons and some of the sugar protons of the uracil. Confirmation from other information is also helpful, such as long-range NOE contacts with neighboring bases, or other knowledge of the structure (which could come from NMR data), or the use of scalar coupling experiments (see Section 1.4.3.3). Evidence for a reversed Hoogsteen base pair was found in natural transfer RNA (yeast tRNA^{Phe}), where the N3H of a uracil showed an NOE cross-peak with the C8H of an adenine [239]. The presence

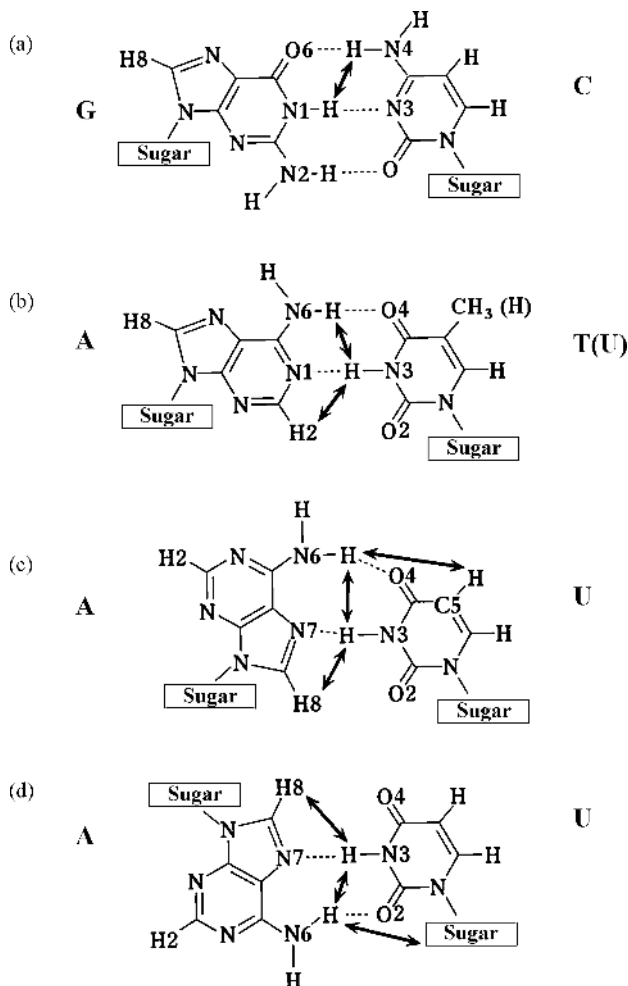


FIGURE 1.55. Examples of common base-pairing schemes, with hydrogen bonding indicated by dotted lines and potential interbase NOEs (dipole couplings) indicated by arrows. The longer arrows [in (c) and (d)] represent NOEs that are generally expected to be weak (4–5 Å), or perhaps not observable. However, the interresidue distances are shown for illustrative purposes only. (a) Watson–Crick-type G–C base pair; (b) Watson–Crick A–T (A–U) base pair; (c) Hoogsteen A–U base pair; (d) reverse Hoogsteen A–U base pair. Many other base-pairing schemes are also possible.

of this reversed Hoogsteen base pair was confirmed by labeling the C8H of adenine with ^2H (which is not observed in ^1H NMR), resulting in the disappearance of the NOE.

The presence or absence of Hoogsteen and reversed Hoogsteen base pairs may also be determined indirectly. One can sometimes examine the conformations of the

glycosidic torsion angles, as the *syn* or *anti* conformations are significantly different (see Section 1.1.8 and Fig. 1.15). If a nucleoside is in a Hoogsteen base pair, its sugar will be in the *syn* conformation, where the interproton distance between C1'H and C8H is very short (2.2 Å), causing a particularly strong NOE cross-peak. Interestingly, this method has been used more commonly to disprove the presence of a Hoogsteen base pair [240].

1.4.4.3. Structure Elucidation from NOE Data There are several ways to generate interproton distances from NOE data, but it is critical to do it accurately. Since the final molecular structure is constructed from the sum of many short distances, even a small systematic error can become amplified when calculating long distances, resulting in a distorted structure. In addition, the protons in nucleic acids are close together, especially in their sugar rings, resulting in interproton NOE cross-peaks with intensities that can be significantly different from those expected. This happens because the magnetization is transferred through spin diffusion networks to other nearby protons. Nucleic acids analysis therefore requires the careful integration of the intensity of each NOE cross-peak.

One of the more accurate methods for calculating interproton NOE distances in nucleic acids is the *complete relaxation matrix analysis*, which determines interproton distances that are independent of an initially proposed structure [241]. This method does not calculate a simple NOE distance from each separate cross-peak, but instead uses general atomic connectivity information within the molecule to determine and analyze the spin diffusion network. A set of model-independent interproton distances (and error estimates) are then iteratively calculated. For this type of analysis, the NOE signal arising from each proton pair is individually integrated and used as input data. Although preliminary structural information is used in this calculation, the final set of interproton distances is independent of the conformations of that structure.

Another approach is the *distance geometry* (DG) algorithm, in which atom placement is based on the geometric network of competing pairs of distances [242]. This method is quite common in protein structure determination, but its use with nucleic acids requires special attention, as nucleic acids (especially DNA) have fewer long-range interresidue NOE peaks. In addition, DG does not normally consider the effects of spin diffusion on the NOE intensities in nucleic acids. This method can be easier to apply than the complete relaxation method, and is particularly useful for the structure determination of DNA and RNA complexes with proteins, drugs, or other ligands, such as in analysis of the complex formed by d(CGCGAATTCGCG)₂ and the ligand, berenil [=1,3-bis(4'-amidinophenyl)triazene] [243].

Regardless of the method by which the set of interproton distances is calculated, molecular modeling is then used to turn that information into an experimentally reasonable structure. The final structures are then verified by comparing them to the experimental data. Details on molecular modeling may be found in Section 1.7.

As it is often difficult to avoid systematic errors when using NOE intensities for model-building purposes, NMR spectroscopists do not rely solely on NOE information, but also use *J*-coupling analyses to calculate pseudorotation and backbone dihedral torsion angles by means of equations similar to the Karplus relationship

(see the preceding section). Since NOE information is limited to interproton distances shorter than 5 Å, important longer-range structural information can sometimes be missed. The technique of residual dipolar coupling analysis (Section 1.4.5.) is complementary to NOE analysis as it can determine interproton directionality vectors (Section 1.4.5). It has a distance cutoff of ~ 8 Å, and hence can provide other essential structural information for DNA.

1.4.4.4. Transfer NOE, Saturation Transfer, and Related Methods The use of NOE data is not restricted to interproton distances within the same molecule, as any proton within 5 Å will be detected, allowing the determination of intermolecular distances between bound ligand molecules (drugs) and DNA. Direct observation of such intermolecular NOEs is usually the first approach attempted, but is often difficult to observe them. It has the disadvantage of working only with tightly bound ligands, for which one often observes two very clearly defined sets of resonances for the bound and unbound states. (The bound ligands in such a system are said to be in “slow exchange” with those that are unbound.) Another disadvantage is that the decreased tumbling rate of macromolecules often leads to broad peaks for both the DNA and the drug (which “goes along for the ride”). To circumvent these problems, several techniques have been designed to utilize the various effects of NOE on a spectrum. These are especially designed for use in drug discovery programs for screening a large pool of ligands to find those that bind to a target receptor molecule, but can be also used to obtain some structural information [244,245].

One important technique is the transfer NOE (trNOE) experiment, which is used to determine whether small ligands can bind to large molecules ($>10,000$ Da), but it works only for drugs that bind weakly ($k_{\text{dissociation}} \sim 10^{-6} - 10^{-3}$ M). Small, unbound ligand molecules tumble rapidly in solution and have small, positive, intra-ligand NOEs. Large macromolecules, like DNA, tumble slowly and have large, negative NOEs. However, when ligands bind to a large molecule, they acquire the tumbling characteristics (defined by the *correlation time* τ_c ; see Appendix A.2.3.1) of that larger receptor molecule, resulting in large, negative NOEs. Thus, bound and unbound ligands can be distinguished by examining the sign and magnitude of the NOE peaks. Typically, the NOEs of an unbound ligand require more than 4 times as much time to reach their maxima, compared to a bound ligand. Hence, by monitoring the signs of *intra-ligand* NOEs, one can determine whether a ligand is bound to a macromolecule. These trNOEs are indicative of the conformation of the bound ligand. The structural conformation of the bound ligand can be determined via experiments such as the 3D TOCSY-trNOE. Intramolecular trNOEs determine the ligand conformation in the bound state, while intermolecular trNOEs can indicate the orientation of the ligands in the binding site [246].

A similar technique is the saturation transfer difference (STD-NOE) experiment. In this method it is not the sign of the intra-ligand NOEs that is monitored, but the transference of magnetization between the molecules. When a macromolecular resonance is selectively irradiated, the spin diffusion networks enable the magnetization to be rapidly dispersed throughout the macromolecule and eventually transferred to any bound ligands (drugs). Ligands that are not tightly bound to the DNA can

dissociate while still carrying the transferred NOE magnetization. Since large molecules, and any bound ligands, have short T_2 relaxation rates (see Fig. 1.48 and Section 1A.2.3.2), they typically have broad, poorly resolved peaks. However, small, unbound, molecules have longer T_2 relaxation rates and narrower peaks. Thus, one can use a relaxation-time “filter” to selectively remove signals from the large complex on the basis of their shorter T_2 values. The STD analysis is performed by acquiring two spectra, one in which the macromolecule has been irradiated, and one in which it is irradiated a great distance from any resonances (effectively, it is unirradiated). The two spectra are subtracted from each other, and the enhanced magnetization is then detected through the unbound ligand molecule, which has a much longer relaxation time [247]. In these experiments a greater than 10-fold excess of the drug is used, which ensures that nearly all of the large receptor molecules (DNA) are involved in binding to the ligand. As this dissociation constant can be in the range of nM to mM, STD-NOE can cover a wider range than trNOE. This technique may be combined with any of the other standard 2D NMR experiments.

1.4.5. Residual Dipolar Coupling

Owing to its short-range potential, NOE information is capable of accurately describing local conformational parameters such as sugar pucker and glycosidic torsion angles. However, NOE is less useful for determining the more macroscopic parameters, such as backbone torsion angles and helical parameters, and detecting an overall bend in the molecule. Even with the addition of J -coupling constraints, reliable global conformations are difficult to obtain.

As noted in Sections 1.4.1 and A.2.3.1), dipolar relaxation arises from the through-space interaction (coupling) of nuclear magnetic dipoles. It is the major source of longitudinal relaxation (Fig. 1.48) measured by T_1 and is the origin of the NOE. Dipolar couplings, such as those of ^1H - ^1H , ^1H - ^{13}C , and ^1H - ^{15}N , are a function of both distance and angular orientation. In (non-spinning) solids the angular orientations are fixed with respect to the direction of the magnetic field $\mathbf{B}_0$, producing very strong dipolar couplings. The resulting relaxation rate is so fast that the NMR spectrum is dominated by very broad and diffuse lines. The effect of this dipolar interaction can extend to many angstroms. In solution, the rapid isotropic tumbling of the molecules causes these dipolar couplings to average out to zero, producing no observable splittings, and the resulting spectrum is comparatively simple, with typically sharp and narrow lines. By adding a liquid crystalline cosolute (see discussion below) it is possible to inhibit this free tumbling and induce a slight orientational preference and partial alignment ($\sim 0.3\%$) of the molecule. This anisotropic rotational diffusion and partial alignment of the molecule results in *residual* dipolar couplings (RDCs) that are weaker than in solids, but stronger than in liquids. There is a growing list of cosolutes that have been used as alignment media. Some alignment media form lipid bicelles, which are disk-shaped particles that typically have a diameter of several hundred angstroms and a thickness of ~ 40 Å. As these lipids are diamagnetic, the bicelles orient with their normal orthogonal to the magnetic field. In addition to bicelles, some other

cosolutes suitable for aqueous solutions are filamentous phage and purple membranes. By varying the concentrations of these media, these residual couplings can be tuned to 10–50 Hz. This is optimal for measuring single bond dipolar couplings, especially those of ^1H - ^{13}C and ^1H - ^{15}N . These bonds are the most commonly studied as they are easily interpreted. Since the interatomic bond distances are essentially fixed, the strength of the single bond residual dipolar coupling will be dependent only on their angular orientations. This allows the construction of a matrix consisting of the orientations of the bonds (the bond vectors). It has the important advantage of enabling the determination of long-range orientational restraints. These constraints are fundamentally different from those obtained by the local NOE and spin coupling constraints, and provide additional data for the structural analysis. From this, in combination with connectivity information, a high-resolution, stereospecific, 3D structure can be created [249].

As an example of the utility of this approach, the solution structure of Dickerson's dodecamer (Section 1.1.5) $\text{d}(\text{CGCGAATTCGCG})_2$ was determined in an aqueous liquid crystalline medium containing 5% w/v (mg/mL) bicelles [249]. This allowed the determination of 198 single-bond dipolar coupling restraints (^{13}C - ^1H and 10 ^{15}N - ^1H), which were in addition to the more standard restraints of 200 approximate ^1H - ^1H dipolar coupling restraints and 162 structurally meaningful NOE restraints. The use of more restraints improves the final structure, producing a DNA duplex that is a highly regular B form, with no significant bends or kinks, and having the expected $\text{C2'}/\text{C1'}$ - exo sugar pseudorotation angles P . These angles, as well as the glycosyl torsion angles and the base-pair propeller twists, were all tightly determined.

When this same dodecamer was later reexamined, an even larger set of residual dipolar couplings were used [250]. This study used the same heteronuclear ^{13}C - ^1H and ^{15}N - ^1H and qualitative homonuclear ^1H - ^1H dipolar couplings that were previously measured in the bicelle medium. Then, an additional 300 quantitative ^1H - ^1H and 22 ^{31}P - ^1H dipolar restraints, together with 22 ^{31}P chemical shift anisotropy restraints, were obtained in a liquid crystalline Pf1 (bacteriophage) medium. The ^{31}P - ^1H dipolar couplings, $\text{H3'}/\text{C3'}/\text{O-}^{31}\text{P}$ spin couplings, and ^{31}P chemical shift anisotropy data were used to restrain the phosphodiester backbone dihedral angles. The final structure represents a quite regular B-form helix with a modest bend of $\sim 10^\circ$, which is essentially independent of whether electrostatic terms are used in the calculation. It was indicated, however, that the dipolar coupling data cannot be fit by a single structure, but are compatible with the presence of a rapid equilibrium between $\text{C2'}/\text{endo}$ and $\text{C3'}/\text{endo}$ deoxyribose puckers. The $\text{C2'}/\text{H2'}/\text{H2''}$ dipolar couplings in B-form DNA were particularly sensitive to sugar pucker and yielded the largest discrepancies when fit to a single structure.

1.5. ELECTRON SPIN RESONANCE SPECTRA

Electron spin resonance (ESR), or electron paramagnetic resonance (EPR), has been the subject of many reviews and books, covering both its theory and its application to

biomolecular systems [251–253]. It is a technique that relies on the same fundamental equations of physics as NMR, except that the spins of electrons (given by the quantum number, $S = \frac{1}{2}$) are excited, rather than the spins of the nuclei. In both NMR and ESR, the sample is placed in a high magnetic field, although the magnetic field strength needed for ESR is considerably less than that for NMR. Since the resonance frequency for an unpaired electron is ~ 1000 times greater than that used for nuclei, ESR requires microwave irradiation to alter the spin populations of electrons, whereas NMR uses radiofrequency.

As with NMR, the ESR signal is influenced by factors such as longitudinal (T_1) and transverse (T_2) relaxation. When an unpaired electron interacts with a nucleus having a spin quantum number of I , it produces a “nuclear hyperfine splitting” of the ESR signal with $(2I + 1)$ lines. Hence, the interaction of an electron with an ^{14}N nucleus ($I = 1$) results in a spectrum with three lines. Unlike NMR spectra, ESR spectra are typically displayed as the first derivative of the absorption spectrum. This enables careful measurement of both the hyperfine splitting constants and the lineshape [254].

The spins of electrons can be detected as an ESR signal only for molecules, ions, or atoms that have unpaired electrons—especially free radicals or transition metal complexes. Since the majority of compounds do not have unpaired electrons under normal conditions, they do not produce an ESR signal. As a result, ESR applications have not been as widespread as those of NMR. However, many proteins and some drug-ligand complexes require the presence of paramagnetic metal ions to be biologically active. In such cases, ESR is a particularly effective way to obtain mechanistic and structural information (Section 5.1.2).

It is also common to modify a compound of interest by adding one or more radicals to it. This is called a *spin label*, such as the nitroxyl radical (see Section 6.2.2, for example). The hyperfine splitting of the ESR signal is sensitive to both the orientation and polarity of the environment of the label; hence it can provide important information about the structure and dynamics of the biomolecular system. More recently, it has become possible to spin-label DNA and RNA, which has greatly aided the study of nucleic acids by ESR. These labels must be used with care, however, as their presence may alter the structure of the complex in such a way that it no longer represents the one that is biological relevant.

1.6. X-RAY CRYSTALLOGRAPHY

This technique examines the transmission and scattering (diffraction) of X rays by a crystal and is used to determine the three-dimensional atomic coordinates of a molecular structure. Crystals can be considered to be infinitely repeating array of identical copies of the molecule of interest, and hence they can produce a scattering signal that is strong enough to be detected. This is not true of single molecules, which lack sufficient scattering power. The photons of X-ray radiation do not interact with the atomic nuclei themselves, but instead interact with the electrons surrounding the atoms. When the photons strike the electron cloud,

they are absorbed and instantaneously re-emitted via a Rayleigh scattering process (Section 1.3.2). Diffraction results from the summation of all the scattered waves of a given wavelength. The resulting diffraction pattern is then analyzed using Bragg's law to give the spacing between atoms or ions in a crystal lattice. This enables the generation of a 3D electron density map, indicating the positions of the heavy atoms (hydrogens are not observed). Using computer modeling approaches, the macromolecular structure is fit to the map.

This brief overview is intended only to provide the most essential background information on this important and specialized technique. The reader is referred to many excellent references on X-ray diffraction methods [255,256] and their applications to DNA, nucleic acids, and drug interactions [10,17,49,257,258].

1.6.1. Bragg Diffraction

The wavelengths associated with X-ray electromagnetic radiation are very short and are of the same order of magnitude as the distance between atoms in crystals ($\sim 1.5 \text{ \AA}$). In a crystal, the atoms are arranged in planes, with each plane stacked on top of another. Bragg determined that the diffraction of X ray by atoms in a crystal can be treated mathematically as if the crystal were a reflection grating [259]. When a monochromatic beam of X ray strikes the stacked planes of a crystal, each plane reflects some of the radiation, while the rest is transmitted through to the next plane, which will also reflect and transmit some of the rays (Fig. 1.56a). The waves of these reflections can be either in-phase with each other, resulting in constructive interference (additive enhancement of the signal), or out-of-phase, resulting in destructive interference (cancellation of the signal). The Bragg equation describes the necessary conditions for constructive interference by relating the wavelength of the X ray (λ) to the distance between the atomic planes in the crystal (d ; the lattice spacing), and the angle between the incident beam and the scattering planes (θ) called the "grazing" or "glancing" angle. The one dimensional form of the Bragg equation is given by

$$n\lambda = 2d \sin \theta \quad (1.11)$$

where n is an integer representing the number of the plane from which the X ray is reflected. The rays striking the lower planes must necessarily travel a greater distance than those striking the top most plane (Fig. 1.56a). When the difference in the distance travelled is equal to an integral number of X-ray wavelengths ($n\lambda$), the emerging beams will be in phase and will reinforce each other. A reflection is called "first-order" when $n = 1$; when $n = 2$, it is called "second-order," and so on. If the X ray strike the crystal at angles other than the grazing angle θ , then the extra distance will not be an integral number of wavelengths, resulting in destructive interference, not reinforcement. In this case, the intensity of the reflected X ray is weakened. Thus, constructive interference occurs only at certain geometries, and when it does, it produces a set of spots (reflections) in the diffraction pattern. Bragg's equation allows the determination of the distance d between the planes of the atoms, if the wavelength of the irradiation is known. The crystal is slowly rotated through a range of θ in order to obtain

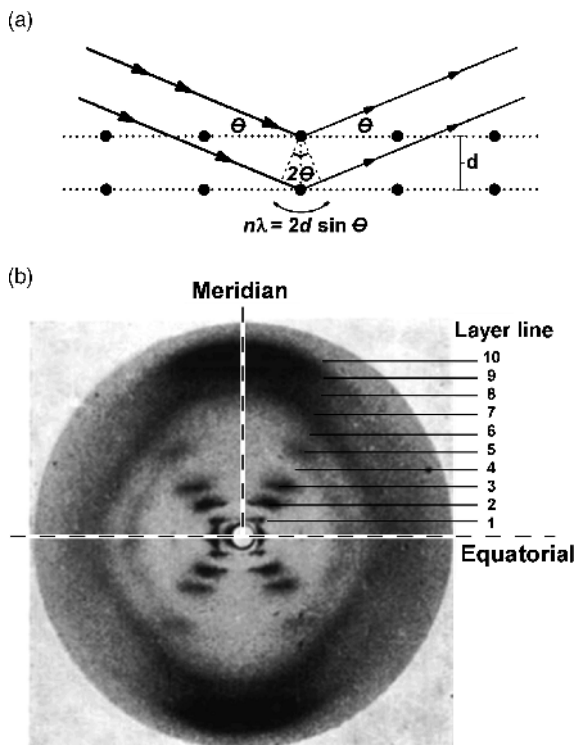


FIGURE 1.56. (a) The scattering of monochromatic light from an array of atoms in a crystal, as described by Bragg's equation. The angles and distances displayed were arbitrarily chosen for illustrative purposes only. (b) Rosalind Franklin's famous photograph of the X-ray fiber diffraction of B-DNA [3], annotated to indicate the layer lines and necessary salient features. The diffraction pattern is a reciprocal lattice image of the DNA fiber, with the distances inversely proportional to those in the fiber.

reflections from as many Bragg planes as possible, and hence increase the number of observed interference spots.

From this information, an electron density map can be generated. The resolution of this map is determined by the quality of the crystal, the X-ray wavelengths used, and the scattering angles observed. The resolution is usually expressed in terms of the interplanar spacings d . At a resolution of $2.5 \text{ \AA}$ the DNA bases begin to be discernable, at $1.5 \text{ \AA}$ they are clearly seen, and by $0.8 \text{ \AA}$ the electron densities of individual atoms are well defined. The highest resolution is usually possible only in crystals of small molecules.

1.6.2. Diffraction of Helical DNA

In order to obtain an accurate three-dimensional model of a macromolecule, it is necessary to obtain atomic resolution diffraction images. Unfortunately, these are

often difficult to obtain because of imperfections in their crystal structure, such as those that arise from impurities, multiple stable conformations, or an inherent asymmetry of the crystal. This is particularly true of polymeric, double-helix DNA, which is more difficult to crystallize than more globular biomolecules, such as RNA, or DNA bound to proteins. (For comparison, consider the great strides that have been made in the X-ray crystal analysis of the ribosome [260–262].) Instead, polymeric DNA has been generally studied as fibers which are drawn from viscous gels, and is referred to as *fiber diffraction*.

A DNA fiber is a continuous helix that can be considered as kind of a one-dimensional crystal. Figure 1.56b shows the now-famous fiber diffraction photograph of B-DNA acquired by Franklin [3]. The diffraction from a helix is scattered perpendicular to the helical axis in a series of planes that appear as lines (layer lines). Bragg spots are not seen because the sample consists of a bundle of helices that are randomly oriented around the helical axis. There is a reciprocal relationship between the distances on the photograph and the distances in real space. The positions of the first and strongest peaks in each layer line describe an X pattern, called the *helical cross*. The angle between the arms of the X is inversely related to the helical radius ($\sim 10 \text{ \AA}$), and the separation of the layer lines is the reciprocal of the pitch of the helix (P), indicating that a complete helical turn is $\sim 34 \text{ \AA}$. The symmetric placement of the spots along the layer lines demonstrates that the helix has a regular twist. The missing spot at the fourth layer line arises because it is here that the two strands of the double helix cross and cancel each other out. The intense reflection on the meridian at the 10th layer line occurs at a distance of $\sim 3.4 \text{ \AA}$, indicating that the helix contains ~ 10 nucleotides per turn, and that the bases are stacked perpendicular to the helical axis with a separation of $\sim 3.4 \text{ \AA}$. The intensity of this reflection results from the fact that the sample was composed of many DNA strands with imperfect alignment and significant rotational disorder. When the fibers are tilted, the reflection from the base pairs is increased, thereby producing the strong meridian reflection on the 10th layer line.

In order to obtain a high-resolution X-ray structure, significant data analysis is required. The intensity of each diffraction spot is proportional to the square of the *structure factor* amplitude, which is a complex number containing information related to both the amplitude and phase of the reflected wave. Once this “phase problem” is solved, an electron density map may be constructed, and from this a molecular model may be built.

1.7. MOLECULAR MODELING AND MOLECULAR MECHANICS

1.7.1. Molecular Modeling

Molecular modeling is a collective term referring to many different theoretical and computational approaches used to calculate molecular structure, molecular dynamics, and molecular properties such as reaction intermediates and thermodynamic analyses. Molecular modeling has become ubiquitous in biomacromolecular

structure determination as there are very few ways to coalesce a large amount of structural data into a 3D structure. It can be a purely theoretical pursuit, but it is also extensively used as a tool in both NMR and X-ray crystallography, where the results are driven by the experimental data. X-ray structure determination is generally more straightforward than it is for NMR because of the intrinsic natures of their experimental data. For this reason, much of what follows pertains to methods that are more often used for NMR.

All the methods used in molecular modeling rely on a set of equations and a large set of parameters, which together compose a *semiempirical forcefield*. The equations attempt to describe mathematically the general motion and interaction energies of atoms. The parameters define the specific characteristics of the atoms and their bonds, and their values are derived from a combination of the results of quantum mechanical (QM) calculations and experimental studies. Much of the underlying experimental and QM data on which these parameters are based are determined from gas-phase or *in vacuo* studies of much smaller, more simple molecules. A brief description of the basic theories of molecular modeling and mechanics can be found in Appendix A.3. For detailed descriptions of the methods and fundamentals of molecular modeling, the interested reader should consult Refs. 263–267.

1.7.2. Modeling of Experimental Data

1.7.2.1. Distance Geometry Unlike most modeling techniques, the distance geometry approach, which is often used in NMR structure refinement (Section 1.4.4), does not use energy functions. Instead, it describes atoms as rigid-sphere potentials (i.e., “solid balls”) and determines their relative three-dimensional positions in a molecule by using connectivity information (bonds) combined with NOE interproton distances. There are also many programs used for crystallographic structure refinement that do not include energy functions. This makes model building more simple, and much faster. Since they are guided by experimental information, the resulting structures are usually very nearly the same as those that have been derived using a combination of energy functions and experimental information. This is not always true, however, and the comparisons can be enlightening.

1.7.2.2. Molecular Dynamics Molecular modeling can also be used to simulate the movement of a molecule by a process called *molecular dynamics* (MD). This technique is often used to model the movement and flexibility of a molecule. It is also frequently used in NMR structure determination as a means of molding an initial structure into one that is experimentally realistic. This is done by subjecting the initial molecular structure to a process called *simulated annealing*, which is very similar to DNA melting curve analysis (Section 1.2.1). The kinetic energy of the atoms in the model structure is increased, such as by raising the temperature (often to 400–600 K), and then, after a short time, it is slowly lowered back to 275 K, where it is allowed to equilibrate for further period of time. This kinetic energy increase allows the molecule

to explore more conformations by allowing bond vibrations and rotations to move more freely. As it “cools back down,” the potential energy surface of the molecule guides it toward a more realistic structure with a relatively lower total energy. The structure is usually then subjected to an *energy minimization* procedure to remove any structural anomalies and produce a lowest-energy structure. During this process an additional component to the potential energy may be imposed on the molecule to represent experimental information. The experimental information usually takes the form of distance or dihedral angle constraints, and functions like tethers or springs that pull or twist the atoms and bonds toward a better fit with the experimental data. These are often derived from NMR NOE and *J*-coupling analyses. When used in this way, this procedure is called *restrained molecular dynamics* (rMD).

Unlike energy minimization, rMD does not result in only one lowest-energy structure, but rather in a family of related structures that together fit the data better than does any single structure alone. This family is usually represented by those structures that are most similar. This is determined by calculating the RMSD (root mean square deviation) between the models. Low RMSD values indicate that the structure calculations have nearly converged to a single structure. The best structures have RMSD values of about <1 Å for the heavy atoms (i.e., excluding hydrogen). If there are large motions in the molecule, such as in loop regions (Section 1.1.11), or if several different conformations exist in solution, the overall RMSD values will be larger. When possible, these are subdivided and analyzed separately. Such situations are often better handled by imposing the experimental NMR information in a time-averaged manner over the course of the simulation (called “MD-tar”) [268], or by performing an *ensemble-average* MD simulation (using multiple copies of the molecule) [269–275] and calculating an averaged RMSD. An alternative approach, which can be combined with the previous methods, is to rigorously analyze the simulations using statistical methods to identify families of structurally related conformers and calculate their relative populations according to the percentages needed to best fit the experimental data [216,275]. This has the advantage of reducing the number of representative structures to an experimentally relevant minimum, and provides a more easily visualized picture of the equilibrium mixture.

NMR is usually used to study macromolecules in solution, which often contain localized regions that undergo greater dynamical motion. When a portion of a molecule is moving, well-defined interatomic distances cannot (and perhaps should not) be imposed on that part of the final structure, resulting in regions that have less stringent structural constraints. This is not the case with X-ray crystallography, where molecular motions are rarely substantial. Such internal motions result in a family of multiple structures that are usually overlaid and displayed in a manner that reveals those regions having greater and lesser structural information. When MD and NMR relaxation experiments are used in combination, one can often obtain a clear picture of the true range of motions existing in solution. This is important, as it is often correlated to changes in molecular conformation upon drug binding.

REFERENCES

1. Watson, J. D. and Crick, F. H. C., "Molecular structure for nucleic acids: A structure for deoxyribose nucleic acid," *Nature* **171**:737–738 (1953).
2. Wilkins, M. H. F., Stokes, A. R., and Wilson, H. R., "Structure of nucleic acids: Molecular structure of deoxypentose nucleic acids," *Nature* **171**:738–740 (1953).
3. Franklin, R. E. and Gosling, R., "Molecular configuration in sodium thymonucleate," *Nature* **171**:740–741 (1953).
4. Rigandy, J. and Klesney, S. P. eds., *International Union of Pure and Applied Chemistry. Nomenclature of Organic Chemistry*, Pergamon Press, Oxford, 1979.
5. Taylor, R. and Kennard, O., "The molecular structures of nucleosides and nucleotides. 1. The influence of protonation on the geometries of nucleic acid constituents," *J. Mol. Struct.* **78**:1–28 (1982).
6. Geise, H. J., Adams, W. J., and Bartell, L. S., "Electron diffraction study of gaseous tetrahydrofuran," *Tetrahedron* **25**:3045–3052 (1969).
7. Almenningen, A., Seip, H. M., and Willardsen, T., "Studies on molecules with five-membered rings. II. An electron diffraction investigation of gaseous tetrahydrofuran," *Acta Chem. Scand.* **23**:2748–2754 (1969).
8. Engelholm, G. G., Gwinn, W. D., and Harris, D. O., "Ring puckering in five-membered rings," *J. Chem. Phys.* **50**:2446–2457 (1969).
9. Cremer, D. and Pople, J. A., "Molecular orbital theory of the electronic structure of organic compounds XXIII. Pseudorotation in saturated five-membered ring compounds," *J. Am. Chem. Soc.* **97**:1358–1367 (1975).
10. Saenger, W., *Principles of Nucleic Acid Structure*, Springer-Verlag, Heidelberg, 1984.
11. Langridge, R., Wilson, H. R., Hooper, C. W., Wilkins, M. H. F., and Hamilton, L. D., "The molecular configuration of deoxyribonucleic acid I. X-ray diffraction study of a crystal-line form of the lithium salt," *J. Mol. Biol.* **2**:19–37 (1960).
12. Langridge, R., Marvin, D. A., Seeds, W. E., Wilson, H. R., Hooper, C. W., Wilkins, M. H. F., and Hamilton, L. D., "The molecular configuration of deoxyribonucleic acid. II. Molecular models and their fourier transforms," *J. Mol. Biol.* **2**:38–64 (1960).
13. Fuller, W., Wilkins, M. H. F., Wilson, H. R., and Hamilton, L. D., "The molecular configuration of deoxyribonucleic acid. IV. X-ray diffraction study of the A form," *J. Mol. Biol.* **12**:60–80 (1965).
14. Arnott, S., Dover, S. D., and Wonacott, A. J., "Least-squares refinement of the crystal and molecular structures of DNA and RNA from X-ray data and standard bond lengths and angles," *Acta. Cryst.*, **B25**:2192–2206 (1969).
15. Arnott, S. and Hukins, D. W., "Optimized parameters for A DNA and B-DNA," *Biochem. Biophys. Res. Commun.* **47**:1504–1510 (1972).
16. Arnott, S., Chandrasekaran, R., Puigjamer, L. C., and Walker, J. K., in eds., Pullman, B. and Jortner, J., Reiter, D., *Nucleic Acids: The Vector of Life*, 1983, pp. 17–31.
17. Dickerson, R. E. and Drew, H. R., "Structure of a B-DNA dodecamer. II. Influence of base sequence on helix structure," *J. Mol. Biol.* **149**:761–786 (1981).
18. Shui, X., McFail-Isom, L., Hu, G. G., and Williams, L. D., "The B-DNA dodecamer at high resolution reveals a spine of water on sodium," *Biochemistry* **37**:8341–8355 (1998).

19. Tereshko, V., Minasov, G., and Egli, M., "The Dickerson-Drew B-DNA dodecamer revised at atomic resolution," *J. Am. Chem. Soc.* **121**:470–471 (1999).
20. Nelson, H. C. M., Finch, J. T., Luisi, F., and Klug, A., "The structure of an oligo(dA)•oligo(dT) tract and its biological implications," *Nature* **330**:221–226 (1987).
21. Wing, R. M., Drew, H. R., Takano, T., Broke, C., Tanaka, S., Itakura, K., and Dickerson, R. E., "Crystal structure analysis of a complete turn of B-DNA," *Nature* **287**:755–758 (1980).
22. DiGabriele, A. D., Sanderson, M. R., and Steitz, T. A., "Crystal lattice packing is important in determining the bend of a DNA dodecamer containing an adenine tract," *Proc. Natl. Acad. Sci. USA* **86**:1816–1820 (1989).
23. Yoon, C., Prive, G. G., Goodsell, D. S., and Dickerson, R. E., "Structure of an alternating B-DNA helix and its relationship to A-tract DNA," *Proc. Natl. Acad. Sci. USA* **85**:6332–6336 (1988).
24. Larsen, T. A., Kopka, M. L., and Dickerson, R. E., "Crystal Structure Analysis of the B-DNA Dodecamer CGTGAATTCACG," *Biochemistry* **30**:4443–4449 (1991).
25. Narayana, N., Ginell, S. L., Russu, I. M., and Berman, H. M., "Crystal and molecular structure of a DNA fragment: d(CGTGAATTCACG)," *Biochemistry* **30**:4449–4455 (1991).
26. Grzeskowiak, K., Yanagi, K., Prive, G. G., and Dickerson, R. E., "The structure of B-helical C-G-A-T-C-G-A-T-C-G and comparison with C-C-A-A-C-G-T-T-G-G," *J. Biol. Chem.* **266**:8861–8889 (1991).
27. Prive, G. G., Yanagi, K., and Dickerson, R. E., "Structure of the B-DNA decamer C-C-A-A-C-G-T-T-G-G and comparison with isomorphous decamers," *J. Mol. Biol.* **217**:177–199 (1991).
28. Heinemann, U., and Alings, C., "Crystallographic study of one twin of G/C rich B-DNA," *J. Mol. Biol.* **210**:369–381 (1989).
29. Timsit, Y., Westhof, E., Fuchs, R. P. P., and Moras, D., "Unusual helical packing in crystals of DNA bearing a mutation hot spot," *Nature* **341**:459–462 (1989).
30. Kennard, O. and Hunter, W. N., "Oligonucleotide structure: A decade of results from single crystal X-ray diffraction studies," *Quart. Rev. Biophys.* **22**:327–379 (1989).
31. Franklin, R. E. and Gosling, R. G., "The structure of sodium thymonucleate fibers. I. The influence of water content," *Acta. Cryst.* **6**:673–677 (1953).
32. Drew, H. R. and Dickerson, R. E., "Structure of B-DNA dodecamer. III. Geometry of hydration," *J. Mol. Biol.* **151**:535–556 (1981).
33. Kopka, M. L., Yoon, C., Goodsell, D., Pjura, P., and Dickerson, R. E., "The molecular origin of DNA-drug specificity in netropsin and distamycin," *Proc. Natl. Acad. Sci. USA* **82**:1376–1380 (1985).
34. Arnott, S., "Polynucleotide secondary structures: an historical perspective," in Neidle, S., ed., *Oxford Handbook of Nucleic Acid Structure*, Oxford Univ. Press, 1999, p. 7.
35. Haran, T. E., Shakked, Z., Wang, A. H.-J., and Rich, A., "The crystal structure of d(CCCCGGGG): A new A-form variant with an extended backbone conformation," *J. Biomol. Struct. Dyn.* **5**:199–217 (1987).
36. Heinemann, U., Lauble, R., Frank, R., and Bloekel, H., "Crystal structure analysis of an A-DNA fragment at 1.8 Å resolution: d(GCCCGGGC)," *Nucl. Acids Res.* **15**:9531–9550 (1987).

37. Robinovich, D., Harqawn, T., Eisenstein, M., and Shakked, Z., "Structures of the mismatched duplex d(GGGTGCCC) and one of its Watson-Crick Analogs, d(GGGCGCCC)," *J. Mol. Biol.* **200**:151–161 (1988).
38. Hunter, W. N., D'Estaintot, B. L., and Kennard, O., "Structural variation in d(CTCTAGAG). Implications for protein-DNA interactions," *Biochemistry* **28**:2444–2451 (1989).
39. Courseille, C., Dautant, A., Hospital, M., D'Estaintot, B. L., Precigoux, G., Molko, D., and Teoule, R., "Crystal structure analysis of an A(DNA) octamer," *Acta Cryst.*, **A46**: FC9–FC12 (1990).
40. Clark, G. R., Brown, D. G., Sanderson, M. R., Chwalinski, T., Meidle, S., Veal, J. M., Jones, R. L., Wilson, W. D., Zon, G., Garman, E., and Stuart, D. I., "Crystal and solution structures of the oligonucleotide d(ATGCGCAT)₂: A combined X-ray and NMR study," *Nucl. Acids Res.* **18**:5521–5528 (1990).
41. Lauble, H., Frank, R., Bloeker, H., and Heinemann, U., "Three-dimensional structure of d(GGGATCCC) in the crystalline state," *Nucl. Acids Res.*, **16**:7799–7816 (1988).
42. Eisenstein, M., Frolov, F., Shakked, Z., and Rabinovich, D., "The structure and hydration of the A-DNA fragment d(GGGTACCC) at room temperature and low temperature," *Nucl. Acids Res.* **18**:3185–3144 (1990).
43. McCall, M., Brown, T., Hunter, W. N., and Kennard, O., "The crystal structure of d(G-G-G-G-C-C-C-C): A model for poly(dG)•poly(dC)," *J. Mol. Biol.* **183**:385–395 (1985).
44. McCall, M., Brown, T., Hunter, W. N., and Kennard, O., "The crystal structure of d(GGATGGGAG) forms an essential part of the binding site for transcription factor IIIA," *Nature* **322**:661–664 (1986).
45. Jain, S., and Sundaralingam, M., "Effect of crystal packing environment on conformation of DNA duplex: Molecular structure of the A-DNA octamer d(G-T-G-T-A-C-A-C) in two crystalline forms," *J. Biol. Chem.* **264**:12780–12784 (1989).
46. Frederick, C. A., Quigley, G. J., Teng, M.-K., Coll, M., van der Marel, G. A., van Boom, J. H., Rich, A., and Wang, A. H. J., "Molecular structure of an A-DNA decamer d(ACCGCCGGT)," *Eur. J. Biochem.* **181**:295–307 (1989).
47. IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN), "Abbreviations and symbols for the description of conformations of polynucleotide chains," *Eur. J. Biochem.* **131**:9–15 (1983).
48. Tsuboi, M., Hirakawa, A. Y., and Nakagawa, M., "A Few comments on the internal rotations in DNA," *J. Mol. Struct.* **126**:525–528 (1985).
49. Dickerson, R. E., Drew, H. R., Conner, B. N., Wing, R. M., Fratini, A. V. and Kopka, M. L., "The anatomy of A-, B-, and Z-DNA," *Science* **216**:475–485 (1982).
50. Schneider, B., Neidle, S., and Berman, H. M., "Conformations of the sugar-phosphate backbone in helical DNA crystal structures," *Biopolymers* **42**:113–124 (1997).
51. Wang, A. H.-J., Quigley, G. J., Kolpak, F. J., Crawford, J. L., van Boom, J. H., van der Marel, G., and Rich, A., "Molecular structure of a left-handed double helical DNA fragment at atomic resolution," *Nature* **282**:680–686 (1979).
52. Gessner, R. V., Frederick, G. A., Quigley, G. J., Rich, A., and Wang, A. H.-J., "The molecular structure of left-handed Z-DNA double helix at 1.0 Å atomic resolution," *J. Biol. Chem.* **264**:7921–7935 (1989).

53. Drew, H. R., Takano, T., Tanaka, S., Itakura, K., and Dickerson, R. E., "High-salt d(CpGpCpG), a left-handed Z' DNA double helix," *Nature* **286**:567–573 (1980).
54. Fujii, S., Wang, A. H.-J., Quigley, G. J., Westerink, H., van der Marel, G., van Boom, J. H., and Rich, A., "The octamers d(CGCGCGCG) and d(CGTATGCG) both crystallize as Z-DNA in the same hexagonal lattice," *Biopolymers* **24**:243–250 (1985).
55. Brennan, R. G., and Sundaralingam, M., "Crystallization and preliminary crystallographic studies of the decadeoxyoligonucleotide d(CpGpTpApCpGpTpApCpG)," *J. Mol. Biol.* **181**:561–563 (1985).
56. Leng, M., in Helene, C., ed., *Structure, Dynamics, Interactions and Evolution of Biological Macromolecules*, Reidel, Dordrecht, 1983, pp. 45–53.
57. Jovin, T. M., McIntosh, L. P., Zarling, D. A., Arndt-Jovin, D. J., Robert-Nicoud, M., and van de Sande, J. H., "Probing for and with left-handed DNA: Poly[d(A-br5C)d(G-T)], a member of a new family of Z-forming DNAs," in Pullman, B. and Tortner, J., eds., *Nucleic Acids: Vector of Life*, Reidel, Dordrecht, 1983, pp. 89–99.
58. Ha, S. C., Lowenhaupt, K., Rich, A., Kim, Y.-G., and Kim, K. K., "Crystal structure of a junction between B-DNA and Z-DNA reveals two extruded bases," *Nature* **437**:1183–1186 (2005).
59. Crick, F. H. C., "Linking numbers and nucleosomes," *Proc. Natl. Acad. Sci. USA* **73**:2639–2643 (1976).
60. Fisher, L. M., Kuroda, R., and Sakai, T. T., "Interaction of bleomycin A₂ with deoxyribonucleic acid: DNA unwinding and inhibition of bleomycin-induced DNA breakage by cationic thiazole amides related to bleomycin A₂," *Biochemistry* **24**:3199–3207 (1985).
61. Utsuno, K., and Tsuboi, M., "Degree of DNA unwinding caused by the binding of aclacinomycin A," *Chem. Pharm. Bull.* **45**:1551–1557 (1997).
62. Hunter, W. N., Brown, T., Anand, N. N., and Kennard, O., "Structure of an adenine-cytosine base pair in DNA and its implication for mismatch repair," *Nature* **320**:552–555 (1986).
63. Patel, D. L., Koslowski, S. A., Ikuta, S., and Itakura, K., "Dynamics of DNA duplexes containing internal G:T, G:A, A:C and T:C pairs: Hydrogen exchange at and adjacent to mismatch sites," *Fed. Proc.* **43**:2665–2670 (1984).
64. Leemor, J.-T., Rabinovich, D., Hope, H., Frolow, F., Appella, E., and Sussman, J. L., "The three dimensional structure of a DNA duplex containing looped-out bases," *Nature* **334**:82–85 (1988).
65. Chattopadhyaya, R., Ikuta, S., Grezeskowiak, K., and Dickerson, R. E., "X-ray structure a DNA hairpin molecule," *Nature* **334**:175–179 (1988).
66. Hare, D. R., and Reid, B. R., "Three-dimensional structure of a DNA hairpin in solution: Two-dimensional NMR studies and distant geometry calculations on d(CGCGTTTTCGCG)," *Biochemistry* **25**:5341–5350 (1986).
67. Sinden, R. R., *DNA Structure and Function*, Academic Press, 1994, p. 135.
68. Felsenfeld, G., and Rich, A., "Studies of the formation of two and three-stranded polynucleotides," *Biochim. Biophys. Acta* **26**:457–468 (1957).

69. Mirkin, S. M., Lyamichev, V. I., Drushlyak, K. N., Dobtynin, V. N., Filippov, S. A., and Frank-Kameneskii, M. D., "DNA H form requires a homopurine-homopyrimidine mirror repeat," *Nature* **330**:495–497 (1987).
70. Hanvey, J. C., Shimizu, M., and Wells, R. D., "Intramolecular DNA triplexes in supercoiled plasmids," *Proc. Natl. Acad. Sci. USA* **85**:6292–6296 (1988).
71. Duval-Valentin, G., de Bizemont, T., Takasugi, M., Mergny, J. L., Bisagni, E., and Helene, C., "Triple-helix specific ligands stabilize H-DNA," *J. Mol. Biol.* **247**:847–858 (1995).
72. Sen, D. and Gilbert, W., "Formation of parallel 4-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis," *Nature* **334**:364–366 (1988).
73. Cech, T. R., "DNA biochemistry—G-strings at chromosome ends," *Nature* **322**:777–778 (1988).
74. Sundquist, W. I., and Klug, A., "Telomeric DNA dimerizes by formation of guanine tetrads between hairpin loops," *Nature* **342**:825–829 (1989).
75. Williamson, J. R., Raghuraman, M. K., and Cech, T. R., "Mono-valent cation induced structure of telomeric DNA—the G-quartet model," *Cell* **59**:871–880 (1989).
76. Blackburn, E. H., "Structure and function of telomeres," *Nature* **350**:569 (1991).
77. Murchie, A. I. H. and Lilley, D. M. J., "Retinoblastoma susceptibility genes contain 5' sequences with a high propensity to form guanine-tetrad structures," *Nucl. Acids Res.* **20**:49–53 (1992).
78. Sundquist, W. I. and Heaphy, S., "Evidence for intrastrand quadruplex formation in the dimerization of human immunodeficiency virus-1 genomic RNA," *Proc. Natl. Acad. Sci. USA* **90**:3393–3397 (1990).
79. Hammond-Kosack, M. C., Kilpatrick, M. W., and Docherty, K., "Analysis of DNA-structure in the human insulin gene-linked polymorphic region *in vivo*," *J. Mol. Endocrinol.* **9**:221–225 (1992).
80. Fry, M. and Loeb, L. A., "The fragile-X syndrome D(Cgg)(N) nucleotide repeats form a stable tetrahelical structure," *Proc. Natl. Acad. Sci. USA* **91**:4950–4954 (1994).
81. Kettani, A., Kumar, R. A., and Patel, D. J., "Solution structure of a DNA quadruplex containing the fragile-X syndrome triplet repeat," *J. Mol. Biol.* **254**:638–656 (1995).
82. Darnell, J. C., Jensen, K. B., Jin, P., Brown, V., Warren, S. T., and Darnell, R. B., "Fragile X mental retardation protein targets G quartet mRNAs important for neuronal function," *Cell* **107**:489–499 (2001).
83. Schaeffer, C., Bardoni, B., Mandel, J. L., Ehresmann, B., Ehresmann, C., and Moine, H., "The fragile X mental retardation protein binds specifically to its mRNA *via* a purine quartet motif," *EMBO J.* **20**:4803–4813 (2001).
84. Brown, B. A., Li, Y. Q., Brown, J. C., Hardin, C. C., Roberts, J. F., Pelsue, S. C., and Shultz, L. D., "Isolation and characterization of a monoclonal anti-quadruplex DNA antibody from autoimmune 'viable motheaten' mice," *Biochemistry* **37**:16325–16337 (1998).
85. Pearson, A. M., Rich, A., and Krieger, M., "Polynucleotide binding to macrophage scavenger receptors depends on the formation of base-quartet-stabilized 4-stranded helices," *J. Biol. Chem.* **268**:3546–3554 (1993).
86. Walsh, K. and Gualberto, A., "Myod binds to the guanine tetrad nucleic-acid structure," *Biol. Chem.* **267**:13714–13718 (1992).

87. Fang, G. and Cech, T. R., "Characterization of a G-quartet formation reaction promoted by the beta-subunit of the oxytricha telomere-binding protein," *Biochemistry* **32**:11646–11657 (1993).
88. Wang, K. Y., Krawczyk, S. H., Bischoftberger, N., Swaminathan, S., and Bolton, P. H., "The tertiary structure of a DNA aptamer which binds to and inhibits thrombin determines activity," *Biochemistry* **32**:11285–11292 (1993).
89. Giraldo, R. and Rhodes, D., "The yeast telomere-binding protein Rap1 binds to and promotes the formation of DNA quadruplexes in telomeric DNA," *EMBO J.* **13**:2411–2420 (1994).
90. Marchand, C., Pourquier, P., Laco, G. S., Jing, N. J., and Pommier, Y., "Interaction of human nuclear topoisomerase I with guanosine quartet-forming and guanosine-rich single-stranded DNA and RNA oligonucleotides," *J. Biol. Chem.* **277**:8906–8911 (2002).
91. Hurley, L. H., "Secondary DNA structures as molecular targets for cancer therapeutics," *Biochem. Soc. Trans.* **29**:692–696 (2001).
92. Gellert, M., Lipsett, M. S., and Davies, D. R., "Helix formation by gyanylic acid," *Proc. Natl. Acad. Sci. USA* **48**:2013–2018 (1962).
93. Arnott, S. and Selsing, E., "Structure of polydeoxyguanylic acid and polydeoxycytidulic acid," *J. Mol. Biol.* **108**:551–552 (1974).
94. Thiele, D. and Guschlbauer, W., "Protonated polynucleotide structures. IX. Disproportionation of poly(G)•poly(C) in acid medium," *Biopolymers* **10**:143–157 (1971).
95. Howard, F. B., Frazier, J., and Miles, H. T., "Stable and metastable forms of poly(G).," *Biopolymers* **16**:791–809 (1977).
96. Henderson, E., Hardin, C. C., Walk, S. K., Tinoco, Jr., K., and Blackburn, E. H., "Telomeric DNA oligonucleotides form novel intramolecular structures containing guanine guanine base-pairs," *Cell* **51**:899–908 (1987).
97. Wang, Y. and Patel, D. J., "Solution structure of the oxytricha telomeric repeat d[(G₄(T₄G₄)₃) G-tetraplex," *J. Mol. Biol.* **251**:76–94 (1995).
98. Aboul-ela, F., Murchie, A. I., and Lilley, D. M., "NMR-study of parallel-stranded tetraplex formation by the hexadeoxynucleotide d(TG₄T)," *Nature* **360**:280–282 (1992).
99. Aboul-ela, F., Murchie, A. I., Norman, D. G., and Lilley, D. M., "Solution structure of a parallel-stranded tetraplex formed by d(TG₄T) in the presence of sodium-ions by nuclear-magnetic-resonance," *J. Mol. Biol.* **243**:458–471 (1994).
100. Wang, Y. and Patel, D. J., "Guanine residues in d(T₂AG₃) and d(T₂G₄) form parallel-stranded potassium cation stabilized G-quadruplexes with antiglycosidic torsion angles in solution," *Biochemistry* **31**:8112–8119 (1992).
101. Wang, Y. and Patel, D. J., "Solution structure of the human telomeric repeat d[1.1.12:AG₃(T₂AG₃)₃] G-tetraplex," *Structure* **1**:263–282 (1993).
102. Smith, F. W. and Feigon, J., "Quadruplex structure of oxytricha telomeric DNA oligonucleotides," *Nature* **356**:164–168 (1992).
103. Smith, F. W. and Feigon, J., "Strand orientation in the DNA quadruplex formed from the oxytricha telomere repeat oligonucleotide d(G₄T₄G₄) in solution," *Biochemistry* **32**:8682–8692 (1992).
104. Strahan, G. D., Shafer, R. H., and Keniry, M. A., "Structural-properties of the d(G₃T₄G₃)₂ quadruplex—evidence for sequential *syn-syn* deoxyguanosines," *Nucl. Acids Res.* **22**:5447–5455 (1994).

105. Smith, F. W., Lau, F. W., and Feigon, J., "d(G₃T₄G₃) forms an asymmetric diagonally looped dimeric quadruplex with guanosine 5'-syn-syn-anti and 5'-syn-anti-anti N-glycosidic conformations," *Proc. Natl. Acad. Sci. USA* **91**:10546–10550 (1994).
106. Ross, W. S. and Hardin, C. C., "Ion-induced stabilization of the G-DNA quadruplex—free-energy perturbation studies," *J. Am. Chem. Soc.* **119**:6070–6080 (1994).
107. Guo, Q., Lu, M., and Kallenbach, N. R., "Adenine affects the structure and stability of telomeric sequences," *J. Biol. Chem.* **267**:15293–15300 (1992).
108. Strahan, G. D., Keniry, M. A., and Shafer, R. H., "NMR structure refinement and dynamics of the K⁺-[d(G₃T₄G₃)]₂ quadruplex via particle mesh Ewald molecular dynamics simulations," *Biophys. J.* **75**:968–981 (1998).
109. Smirnov, I. and Shafer, R. H., "Effect of loop sequence and size on DNA aptamer stability," *Biochemistry* **39**:1462–1468 (2000).
110. Gu, J. D. and Leszczynski, J., "Structures and properties of mixed DNA bases tetrads: Nonempirical ab initio HF and DFT studies," *J. Phys. Chem. A* **104**:1898–1904 (2000).
111. Hud, N. V., Smith, F. W., Anet, F. A. L., and Feigon, J., "The selectivity for K⁺ versus Na⁺ in DNA quadruplexes is dominated by relative free energies of hydration: A thermodynamic analysis by H-1 NMR," *Biochemistry* **35**:15383–15390 (1996).
112. Miyoshi, D., Karimata, H., and Sugimoto, N., "Hydration regulates thermodynamics of G-quadruplex formation under molecular crowding conditions," *J. Am. Chem. Soc.* **128**:7957–7963 (2006).
113. Marathias, V. M., Wang, K. Y., Kumar, S., Pham, T. Q., Swaminathan, S., and Bolton, P. H., "Determination of the number and location of the manganese binding sites of DNA quadruplexes in solution by EPR and NMR in the presence and absence of thrombin," *J. Mol. Biol.* **260**:378–394 (1996).
114. Marathias, V. M. and Bolton, P. H., "Determinants of DNA quadruplex structural type: Sequence and potassium binding," *Biochemistry* **38**:4355–4364 (1999).
115. Lee, J. S., "The stability of polypurine tetraplexes in the presence of monovalent and divalent-cations," *Nucl. Acids Res.* **20**:6057–6060 (1990).
116. Hardin, C. C., Watson, T., Corregan, M., and Bailey, C., "Cation-dependent transition between the quadruplex and Watson-Crick hairpin forms of d(CGCG₃GCG)," *Biochemistry* **31**:833–841 (1992).
117. Jin, R., Gaffney, B. L., Wang, C., Jones, R. A., and Breslauer, K., "Thermodynamics and structure of a DNA tetraplex: A spectroscopic and calorimetric study of the tetramolecular complexes of d(TG₃T) and d(TG₃T₂G₃T)," *J. Proc. Natl. Acad. Sci. USA* **89**:8832–8836 (1992).
118. Lu, M., Guo, Q., and Kallenbach, N. R., "Thermodynamics of G-tetraplex formation by telomeric DNAs," *Biochemistry* **31**:2455–2459 (1993).
119. Chen, F.-M., "Sr²⁺ facilitates intermolecular G-quadruplex formation of telomeric sequences," *Biochemistry* **31**:3769–3776 (1992).
120. Xu, Q., Deng, H., and Braunlin, W. H., "Selective localization and rotational immobilization of univalent cations in quadruplex DNA," *Biochemistry* **32**:13130–13137 (1993).
121. Miura, T. and Thomas, G. J., Jr., "Structural polymorphism of telomere DNA—inter-quadruplex and duplex-quadruplex conversions probed by Raman-spectroscopy," *Biochemistry* **33**:7848–7456 (1994).

122. Miura, T., Benevides, J. M., and Thomas, G. J., Jr., "A phase-diagram for sodium and potassium-ion control of polymorphism in telomeric DNA," *J. Mol. Biol.* **248**:233–238 (1995).
123. Deng, H. and Braunlin, W. H., "Kinetics of sodium ion binding to DNA quadruplexes," *J. Mol. Biol.* **255**:476–483 (1996).
124. Töhl, J. and Eimer, W., "Interaction energies and dynamics of alkali and alkaline-earth cations in quadruplex-DNA structures," *J. Mol. Model.* **2**:327–329 (1996).
125. Hardin, C. C., Corregan, M. J., Lieberman, D. V., and Brown, B. A., "Allosteric interactions between DNA strands and monovalent cations in DNA quadruplex assembly: Thermodynamic evidence for three linked association pathways," *Biochemistry* **35**:15383–15390 (1997).
126. Gueron, M., Eisinger, J., and Shulman, R. G., "Excited states of nucleotides and singlet energy transfer in polynucleotides," *J. Chem. Phys.* **47**:4077–4091 (1967).
127. Sueoka, N. and Cheng, T. Y., "Natural occurrence of a deoxyribonucleic acid resembling the deoxyadenylate-deoxythymidylate polymer," *Proc. Natl. Acad. Sci. USA* **48**:1851–1856 (1962).
128. Tinoco, I., Jr., "Hypochromism in polynucleotides," *J. Am. Chem. Soc.* **82**:4785–4790 (1960) and "Additions and corrections—hypochromism in polynucleotides," *J. Am. Chem. Soc.* **83**:5047 (1961).
129. Marmur, J. and Doty, P., "Determination of the base composition of deoxyribonucleic acid from its thermal denaturation temperature," *J. Mol. Biol.* **5**:109–118 (1962).
130. Honnas, P. I. and Steen, H. B., "X-ray- and U.V.-induced excitation of adenine, thymine and related nucleosides and nucleotides in solution at 77°K," *Photochem. Photobiol.* **11**:67–76 (1970).
131. Vigny, P., "Fluorescence des nucleosides et nucleotides a temperature ambiante," *C. R. Acad. Sci.*, **D272**:3206–3209 (1971).
132. Takahashi, S., Nishimura, Y., and Tsuboi, M., "Rate of photochemical protonation and electronic relaxation of excited I,N6-ethenoadenosine in its aqueous solution," *J. Chem. Phys.* **15**:3831–3837 (1981).
133. Nishimura, Y., Yamamoto, T., and Tsuboi, M., "The nature of fluorescence spectra of some minor bases obtained from tRNA and mRNA," *Nucl. Acid Res. (Spec. Publ. (3))* 85–88 (1977).
134. Kasai, H., Ohashi, Z., Harada, F., Nishimura, S., Oppenheimer, N. J., Crain, P. F., Liehr, J. G., von Minden, D. L., and McClosky, J. A., "Structure of the modified nucleoside O isolated from *Escherichia coli* transfer ribonucleic acid. 7-(4,5-cis-Dihydroxy-1-cyclopentene-3-ylaminomethyl)-7-deazaguanosine," *Biochemistry* **14**:4198–4208 (1975).
135. Thiebe, R., Zachau, G., Baczynskyj, L., Biemann, K., and Sonnenbichler, J., "Study on the properties and structure of the modified base Y⁺ of yeast tRNA^{Phe}," *Biochim. Biophys. Acta* **240**:163–169 (1971).
136. Kohwi-Shigematsu, T., Enomoto, T., Yamada, M., Nakanishi, M., and Tsuboi, M., "Exposure of DNA bases induced by the interaction of DNA and calf thymus DNA helix-destabilizing protein," *Proc. Natl. Acad. Sci. USA* **75**:4689–4693 (1978).
137. Sutherland, G. B. B. M., and Tsuboi, M., "The infrared spectrum and molecular configuration of sodium deoxyribonucleate," *Proc. Roy. Soc. A (London)* **239**:446–463 (1957).

138. Thulstrup, E. W. and Michl, J., *Elementary Polarization Spectroscopy*, Wiley, New York, 1989.
139. Rodger, A. and Norden, B., *Circular Dichroism and Linear Dichroism*, Oxford Univ. Press, 1997.
140. Kinoshita, K., Jr., Kawato, S., and Ikegami, A., "A theory of fluorescence polarization decay in membranes," *Biophys. J.* **20**:289–305 (1977).
141. Lipari, G. and Szabo, A., "Effect of libration motion on fluorescence depolarization and nuclear magnetic resonance relaxation in macromolecules and membranes," *Biophys. J.* **30**:489–506 (1980).
142. Norden, B., and Kurucsev, T., "Analyzing DNA complexes by circular and linear dichroism," *J. Mol. Recogn.* **7**:141–155 (1994).
143. Fritzsche, H., "Infrared linear dichroism studies of DNA-drug complexes: Quantitative determination of the drug-induced restriction of the B-A transition," *Nucl. Acids Res.* **22**:787–791 (1994).
144. Fasman, G. D., ed., *Circular Dichroism and the Conformational Analysis of Biomolecules*, Plenum Press, New York, 1996.
145. Bush, C. A. and Brahma, J., "Optical activity of single-stranded oligonucleotides," *J. Chem. Phys.* **46**:79–88 (1967).
146. Matsuzaki, J., Hotoda, H., Sekine, M., Hata, T., Higuchi, S., Nishimura, Y., and Tsuboi, M., "Self-complementary tetradexyribonucleoside triphosphates: convenient chemical preparation and spectroscopic studies in solution," *Tetrahedron* **42**:501–531 (1986).
147. Samejima, T., Hashizume, H., Imahori, K., Fujii, I., and Miura, K., "Optical rotatory dispersion and circular dichroism of rice dwarf virus ribonucleic acid," *J. Mol. Biol.* **34**:39–48 (1968).
148. Allen, F. S., Gray, D. M., Roberts, G. P., and Tinoco, I., Jr., "The ultraviolet circular dichroism of some natural DNAs and an analysis of the spectra for sequence information," *Biopolymers* **11**:853–879 (1972).
149. Pohl, F. M., and Jovin, T. M., "Salt-induced co-operative conformational change of sythetic DNA: Equilibrium and kinetic studies with poly(dG-dC)," *J. Mol. Biol.* **67**:375–396 (1972).
150. Wilson, E. B., Decius, J. C., and Cross, P. C., *Molecular Vibrations*, McGraw-Hill, New York, 1955, Chap. 10.
151. Aida, M., Kaneko, M., Dupuis, M., Ueda, T., Ushizawa, K., Ito, G., Kumakura, A., and Tsuboi, M., "Vibrational modes in thymine molecule from an ab initio MO calculation," *Spectrochim. Acta* **A53**:293–407 (1997).
152. Lee, A., Adamowicz, L., Nowak, M. J., and Lapinski, L., "The infrared spectra of matrix isolated uracil and thymine: An assignment based on new theoretical calculations," *Spectrochim. Acta* **48A**:1385–1395 (1992).
153. Tsuboi, M., Nishimura, Y., Hirakawa, A. Y., and Peticolas, W. L., "Resonance Raman spectroscopy and normal modes of the nucleic acid bases," in Spiro, T. G., ed., *Biological Applications of Raman Spectroscopy, Vol. 2, Resonance Raman Spectra of Polymers and Aromatics*, Wiley, New York, 1987, pp. 109–179.
154. Tsuboi, M., "New approaches to the analysis of vibrations of nucleic acids and their components," in Stepanek, J., Anzenbacher, P., and Sedlacek, B., eds., *Laser Scattering Spectroscopy of Biological Objects*, Vol. 45 in Studies in Physical and Theoretical Chemistry (series), Elsevier, 1987, pp. 351–368.

155. Susi, H., Ard, J. S., and Purcell, J. M., "Vibrational spectra of nucleic acid constituents II. Planar vibrations of cytosine," *Spectrochim. Acta* **29A**:725–733 (1973).
156. Nishimura, Y., and Tsuboi, M., "In-plane vibrational modes of cytosine from *ab initio* MO calculation," *Chem. Phys.* **98**:71–80 (1985).
157. Angell, C. L., "An infrared spectroscopic investigation of nucleic acid constituents," *J. Chem. Soc. (Pt. 1)* 504–515 (1961).
158. Hirakawa, A. Y., Okada, H., Sasagawa, S., and Tsuboi, M., "Infrared and Raman spectra of adenine and its ^{15}N and ^{13}C substitution products," *Spectrochim. Acta* **41A**:209–216 (1985).
159. Tsuboi, M. and Kyogoku, Y., "Infrared spectroscopy of nucleic acid components," in Zorbach, W. W. and Tipson, S., eds., *Synthetic Procedures in Nucleic Acid Chemistry*, Vol. II, Wiley, 1973, Chap. 6.
160. Nishimura, Y., Tsuboi, M., Kato, S., and Morokuma, K., "In-plane vibrational modes of guanine from an *ab initio* MO calculation," *Bull. Chem. Soc. Jpn.* **58**:638–645 (1985).
161. Kyogoku, Y., Lord, R. C., and Rich, A., "Hydrogen bonding specificity of nucleic acid purines and pyrimidines in solution," *Science* **154**:518–520 (1966).
162. Delabar, J. M. and Majoube, M., "Infrared and Raman spectroscopic study of ^{15}N and D-substituted guanines," *Spectrochim. Acta* **34A**:129–140 (1978).
163. Lane, M. and Thomas, G. J., "Kinetics of hydrogen-deuterium exchange in guanosine determined by laser-Raman spectroscopy," *Biochemistry* **18**:3839–3849 (1979).
164. Herzberg, G., *Molecular Spectra and Molecular Structure; Infrared and Raman of Polyatomic Molecules*, Krieger, (reprint edition), New York, 1991.
165. Tsuboi, M., Ueda, T., Ushizawa, K., Sasatake, Y., Ono, A., Kainosho, M., and Ishido, Y., "Assignments of the deoxyribose vibrations: isotopic thymidine," *Bull. Chem. Soc. Jpn.* **67**:1483–1484 (1994).
166. Tsuboi, M., Komatsu, M., Hoshi, J., Kawashima, E., Sekine, T., Ishido, Y., Russel, M. P., Benevides, J. M., and Thomas, G. J., Jr., "Raman and infrared spectra of (2'S)-[2-2H] thymidine: Vibrational coupling between deoxyribosyl and thymine moieties and structural implications," *J. Am. Chem. Soc.* **119**:2025–2032 (1997).
167. Tsuboi, M., Takeuchi, Y., Kawashima, E., Ishido, Y., and Aida, M., "Raman spectrum of [^{12}C]thymidine: Vibrations of its 5'-end atomic group," *Spectrochim. Acta* (Pt. A) **55**:1887–1896 (1999).
168. Toyama, A., Takino, Y., Takeuchi, H., and Harada, I., "Ultraviolet resonance Raman spectra of ribosyl C(1')-deuterated purine nucleosides; evidence of vibrational coupling between purine and ribose rings," *J. Am. Chem. Soc.* **115**:11092–11098 (1993).
169. Nishimura, Y., Tsuboi, M., Sato, T., and Aoki, K., "Conformation-sensitive Raman lines of mononucleotides and their use in a structure analysis of polynucleotides: Guanine and cytosine nucleotides," *J. Mol. Struct.* **146**:123–153 (1986).
170. Tsuboi, M., "Vibrational spectra of phosphite and hypophosphite anions, and the characteristic frequencies of PO_3^{2-} and PO_2^- groups," *J. Am. Chem. Soc.*, **79**:1351–1353 (1957).
171. Tsuboi, M., "Infrared and Raman spectroscopy," in Ts'o, P. O. P., ed., *Basic Principles of Nucleic Acid Chemistry*, Vol. 1, Academic Press, New York, 1974, p. 427.

172. Tsuboi, M., Takahashi, S., and Harada, I., "Infrared and Raman spectra of nucleic acids—vibrations in the base-residues," in Duchesne, J., ed., *Physico-Chemical Properties of Nucleic Acids*, Academic Press, New York, 1973, pp. 91–145.
173. Ueda, T., Ushizawa, K., and Tsuboi, M., "Depolarization of Raman scattering from some nucleotides of RNA and DNA," *Biopolymers* **33**:1791–1803 (1990).
174. Kubasek, W. L., Hudson, B., and Peticolas, W. L., "Ultraviolet resonance Raman excitation profiles of nucleic acid bases with excitation from 200 to 300 nanometers," *Proc. Natl. Acad. Sci. USA* **82**:2369–2373 (1985).
175. Fodor, S. P. A., Rava, R. P., Hays, T. R., and Spiro, T. G., "Ultraviolet resonance spectroscopy of the nucleotides with 266, 240, 218 and 200 nm pulsed laser excitation," *J. Am. Chem. Soc.* **107**:1520–1529 (1985).
176. Shimanouchi, T., Tsuboi, M., and Kyogoku, Y., "Infrared spectra of nucleic acids and related compounds," *Adv. Chem. Phys.* **7**:435–498 (1964).
177. Thomas, G. J., Jr., and Tsuboi, M., "Raman spectroscopy of nucleic acids and their complexes," *Adv. in Biophys. Chem.* **3**:1–70 (1993).
178. Tsuboi, M., "Application of infrared spectroscopy to structure study of nucleic acids," *Appl. Spectrosc. Rev.* **3**:45–90 (1969).
179. Nishimura, Y., Morikawa, K., and Tsuboi, M., "Spectral difference of the A and B forms of deoxyribonucleic acid," *Bull. Chem. Soc. Jpn.* **47**:1043–1044 (1974).
180. Pilet, J. and Brahms, J., "Investigation of DNA structural changes by infrared spectroscopy," *Biopolymers* **12**:387–403 (1973).
181. Nishimura, Y., Morikawa, K., and Tsuboi, M., "Infrared spectra of deoxyribonucleic acids with different base compositions in their D₂O solutions," *Bull. Chem. Soc. Jpn.* **46**:3891–3892 (1973).
182. Flavell, R. A. and Jones, I. G., "Mitochondrial deoxyribonucleic acid from *Tetrahymena pyriformis* and its kinetic complexity," *Biochem. J.* **116**:811–817 (1970).
183. Ulitzur, S., "Rapid determination of DNA base composition by ultraviolet spectroscopy," *Biochim. Biophys. Acta* **272**:1–11 (1972).
184. Sueoka, N. and Cheng, T. Y., "Fractionation of nucleic acids with the methylated albumin column," *J. Mol. Biol.* **4**:161–172 (1962).
185. Thomas, G. J. Jr., Benevides, J. M., Overman, S. A., Ueda, T., Ushizawa, K., Saitoh, M., and Tsuboi, M., "Polarized Raman spectra of oriented fibers of A DNA and B DNA: Anisotropic and isotropic local Raman tensors of base and backbone vibrations," *Biophys. J.* **68**:1073–1088 (1995).
186. Tsuboi, M. and Thomas, G. J., Jr., "Raman scattering tensors in biological molecules and their assembly," *Appl. Spectrosc. Rev.* **32**:263–299 (1997).
187. Erfurth, S. C., Kiser, E. J., and Peticolas, W. J., "Determination of the backbone structure of nucleic acids and nucleic acid oligomers by laser Raman scattering," *Proc. Natl. Acad. Sci. USA* **69**:938–941 (1972).
188. Prescott, B., Steinmetz, W., and Thomas, G. J., Jr., "Characterization of DNA secondary structure by Raman spectroscopy," *Biopolymers* **23**:235–256 (1984).
189. Guan, Y., Eurrey, C. J., and Thomas, G. J., Jr., "Vibrational analysis of nucleic acids I. The phosphodiester group in dimethyl phosphate model compounds: (CH₃O)₂PO₂[−], (CD₃O)₂PO₂[−], and (¹³CH₃O)₂PO₂[−]," *Biophys. J.* **66**:225–235 (1994).

190. Benevides, J. M., Tsuboi, M., Wang, A. H.-J., and Thomas, G. J., Jr., "Local Raman tensors of double-helical DNA in the crystal: A basis for determining DNA residue orientations," *J. Am. Chem. Soc.* **115**:5351–5359 (1993).
191. Bovey, F. A., Jelinski, L., and Mirau, P. A., *Nuclear Magnetic Resonance Spectroscopy*, Academic Press, San Diego, 1988.
192. Claridge, T. D. W., *High Resolution NMR Techniques in Organic Chemistry*, Pergamon Press, Elsevier Science, Amsterdam, 1999.
193. Wehrli, F. W. and Wirthlin, T., *Interpretation of Carbon-13 NMR Spectra*, Heyden, London, 1976.
194. Cavanagh, J., Fairbrother, W. J., Palmer, A. G., III, and Skelton, N. J., *Protein NMR Spectroscopy: Principles and Practice*, Academic Press, San Diego, 1996.
195. Friebolin, H., *Basic One- and Two-Dimensional NMR Spectroscopy* (transl. J. K. Becconsall), Wiley-VCH, Weinheim, Germany, 2005.
196. Berger, S. and Braun, S., 200 and More NMR Experiments. A Practical Course, Wiley-VCH, Weinheim, Germany, 1998.
197. Van de Ven, F. J. and Hilbers, C. W., "Nucleic acids and nuclear magnetic resonance," *Eur. J. Biochem.* **178**:1–38 (1988).
198. Wüthrich, K., *NMR of Proteins and Nucleic Acids*, Wiley-Interscience, New York, 1986.
199. Shafer, R. H. and Brown, S. C., "Nuclear magnetic resonance studies of drug-nucleic acid interactions," in Kallenbach, N. R., ed., *Chemistry & Physics of DNA-Ligand Interactions*, Adenine Press, Schenectady, NY, 1990, pp. 109–142.
200. Zidek, L., Stefl, R., and Sklenar, V., "NMR methodology for the study of nucleic acids," *Curr. Opin. Struct. Biol.* **11**:275–281 (2001).
201. Driscoll, P. C., Esposito, D., and Pfuhl, M., "NMR of nucleic acids and proteins," *Nucl. Magn. Reson.* **29**:340–405 (2000).
202. Varani, G., Aboul-ela, F., and Allain, F. H.-T., "NMR investigation of RNA structure," *Prog. NMR Spectrosc.* **29**:51–127 (1996).
203. Flinders, J. and Dieckmann, T., "NMR spectroscopy of ribonucleic acids," *Prog. NMR Spectrosc.* **48**:137–159 (2006).
204. Harris, R. K., Becker, E. D., Cabral de Menezes, S. M., Goodfellow, R., and Granger, P., "Nuclear spin properties and conventions for chemical shifts (IUPAC Recommendations, 2001)," in *Encyclopedia of Nuclear Magnetic Resonance*, Vol. 9, Wiley, Chichester, UK, 2002, p. 5.
205. Takahashi, S., Nagashima, N., Nishimura, Y., and Tsuboi, M., "Anomalous temperature dependence of the phosphorus-31 nuclear magnetic resonance chemical shift in d(CCGG) and d(CCTAGG) at the junction of the pyrimidine stack followed by the purine stack," *Chem. Pharm. Bull.* **34**:3987–3993 (1986).
206. Takahashi, S., Nagashima, N., Nishimura, Y., and Tsuboi, M., "Nuclear magnetic resonance study on the interaction of aclacinomycin-A with a deoxyribo-hexanucleotide Pentaphosphate d(CCTAGG)₂ in Aqueous Solution," *Chem. Pharm. Bull.* **34**:4494–4499 (1986).
207. Bolton, P. H. and James, T. L., "Fast and slow conformational fluctuations of RNA and DNA. Subnanosecond internal motion correlation times determined by ³¹P NMR," *J. Am. Chem. Soc.* **102**:25–31 (1980).

208. Neumann, J. M., Borrel, J., Thiery, J. M., Guschlbauer, W., and Tran-Dinh, S., "PMR-relaxation and steric computations give unequivocal nucleoside conformations," *Biochim. Biophys. Acta* **479**:427–440 (1977).
209. Biological Magnetic Resonance Bank, Univ. Wisconsin-Madison, Madison, 2006 (<http://www.bmrb.wisc.edu>).
210. Jardetzky, O. and Roberts, G. C. K., *NMR in Molecular Biology*, Academic Press, 1981, p. 188.
211. Karplus, M., "Contact electron-spin coupling of nuclear magnetic moments," *J. Chem. Phys.* **30**:11–15 (1959).
212. Haasnoot, C. A. G., de Leeuw, F. A. A. M., Leeuw, H. P. M., and Altona, C., "The relationship between proton-proton NMR coupling constants and substituent electronegativities. II. Conformational analysis of the sugar ring in nucleosides and nucleotides in solution using a generalized Karplus equation," *Org. Magn. Reson.* **15**:43–52 (1981).
213. Rinkel, L. J. and Altona, C., "Conformational analysis of the deoxyribofuranose ring in DNA by means of sums of proton-proton coupling constants: A graphical method," *J. Biomol. Struct. Dyn.* **4**:621–649 (1987).
214. Gonzalez, C., Stec, W., Kobylanska, A., Hogrefe, R. I., Reynolds, M., and James, T. L., "Structural study of a DNA/RNA hybrid duplex with a chiral phosphorothioate moiety by NMR: Extraction of distance and torsion angle constraints and imino proton exchange rates," *Biochemistry* **33**:11062–11072 (1994).
215. Schmitz, U., Zon, G., and James, T. L., "Deoxyribose conformation in [d(GTATATA)]₂: Evaluation of sugar pucker by simulation of double-quantum-filtered COSY cross-peaks," *Biochemistry* **29**:2357–2368 (1990).
216. Ulyanov, N. B., Schmitz, U., Kumar, A., and James, T. L., "Probability assessment of conformational ensembles: Sugar repuckering in a DNA duplex solution," *Biophys. J.* **68**:13–24 (1995).
217. Ono, A., Tate, S., and Kainosho, M., "Preparation and heteronuclear 2D NMR spectroscopy of a DNA dodecamer containing a thymidine residue with a uniformly ¹³C-labeled deoxyribose ring," *J. Biomol. NMR* **4**:581–586 (1994).
218. Pervushin, K., Ono, A., Fernandez, C., Szyperski, T., Kainosho, M., and Wüthrich, K., "NMR scalar couplings across Watson-Crick base pair hydrogen bonds in DNA observed by transverse relaxation-optimized spectroscopy," *Proc. Natl. Acad. Sci. USA* **95**:14147–14151 (1998).
219. Dingley, A. J. and Grzesiek, S., "Direct observation of hydrogen bonds in nucleic acid base pairs by internucleotide ²J_{NN} couplings," *J. Am. Chem. Soc.* **120**:8293–8297 (1998).
220. Dingley, A. J., Masse, J. E., Peterson, R. D., Barfield, M., Feigon, J., and Grzesiek, S., "Internucleotide scalar couplings across hydrogen bonds in Watson-Crick and Hoogsteen base pairs of a DNA triplex," *J. Am. Chem. Soc.* **121**:6019–6027 (1999).
221. Wöhnert, J., Dingley, A. J., Stoldt, M., Görlach, M., Grzesiek, S., and Brown, L., "Direct identification of NH⋯N hydrogen bonds in non-canonical base pairs of RNA by NMR spectroscopy," *Nucl. Acids Res.* **27**:3104–3110 (1999).
222. Grzesiek, S., Cordier, F., Jaravine, V., and Barfield, M., "Insights into biomolecular hydrogen bonds from hydrogen bond scalar couplings," *Prog. NMR Spectrosc.* **45**:275–300 (2004).

223. Tsuboi, M., Kuriyagawa, F., Matso, K., and Kyogoku, Y., "Nuclear spin coupling and conformation of the P—O—C—H group," *Bull. Chem. Soc. Jpn.* **40**:1813–1818 (1967).
224. Kainosho, M., Nakamura, A., and Tsuboi, M., "Phosphorus—proton spin-spin coupling in the P—O—C—H group: A comparison of cyclic and acyclic systems," *Bull. Chem. Soc. Jpn.* **42**:1713–1718 (1969).
225. Wood, D. J., Mynott, R. J., Hruska, F. E., and Sarma, R. H., "¹H NMR study of the molecular conformation of the exocyclic linkage of 5'-B-nucleotides in solution—correlation between C(4')-C(5') and C(5')-O(5') conformer population and the influence of the base," *FEBS Lett.* **34**:323–326 (1973).
226. Lankhorst, P. P., Haasnoot, C. A. G., Erhelens, C., and Altona, C., "Carbon-13 NMR in conformational analysis of nucleic acid fragments 2. A reparameterization of the Karplus equation for vicinal NMR coupling constants in CCOP and HCOP fragments," *J. Biomol. Struct. Dyn.* **1**:1387–1405 (1984).
227. Roongta, V. A., Jones, C. R., and Gorenstein, D. G., "Effect of distortions in the deoxyribose phosphate backbone conformation of the duplex oligonucleotide dodecamers GT, GG, GA, AC and GU base-pair mis-matches on ³¹P NMR spectra," *Biochemistry* **29**:5245–5258 (1990).
228. Majumdar, A. and Hosur, R. V., "Simulation of 2D NMR spectra for determination of solution conformations of nucleic acids," *Prog. NMR Spectrosc.* **24**:109–158 (1992).
229. Wijmenga, S. S., Mooren, M. W., and Hilbers, C. W., "NMR of nucleic acids: From spectrum to structure," in Roberts, G. C. K., ed., *NMR of Macromolecules: A Practical Approach*, Oxford Univ. Press, 1993, pp. 217–283.
230. Chary, K. V. R., Rastogi, V. K., Govil, G., and Miles, T., "Estimation of ³¹P-¹H and ¹H-¹H vicinal coupling constants along the DNA backbone by 2D HELCO measurements," *J. Chem. Soc. Chem. Commun.* **1994**:241–243 (1994).
231. Ojha, R. P., Dhingra, M. M., Sarma, M. H., Shibata, M., Farrar, M., Turner, C. J., and Sarma, R. H., "DNA bending and sequence-dependent backbone conformation: NMR and computer experiments," *Eur. J. Biochem.* **265**:35–53 (1999).
232. Kim, S. G., Lin, L. J., and Reid, B. R., "Determination of nucleic acid backbone conformation by ¹H NMR," *Biochemistry* **31**:3566–3574 (1992).
233. Ding, D., Gryazonv, S. M., Lloyd, D. H., Chandrasekaran, S., Yao, S., Ratmeyer, J., Pan, Y., and Wilson, W. D., "An oligonucleotide N3'-P5' phosphoramidate duplex forms an A-type helix in solution," *Nucl. Acids Res.* **24**:354–360 (1996).
234. Onoye, H., Suzuki, E., Nishimura, Y., and Tsuboi, M., "Proton magnetic resonance studies of d(TGGGGT) and other oligodeoxyribonucleotides," (in press).
235. Haasnott, C. A. G. and Hilbers, C. W., "Effective water resonance suppression in 1D and 2D-FT-¹H nmr spectroscopy of biopolymers in aqueous solution," *Biopolymers* **22**:1259–1266 (1983).
236. Otting, G., Grutter, R., Leupin, W., Minganti, C., Ganesh, K. N., Sproat, B. S., Gait, M. J., and Wüthrich, K., "Sequential NMR assignments of labile protons in DNA using two-dimensional nuclear-Overhauser-enhancement spectroscopy with three jump-and-return pulse sequences," *Eur. J. Biochem.* **166**:215–220 (1987).
237. Barbault, F., Huynh-Dinh, T., Paoletti, J., and Lancelot, G., "A peculiar DNA structure: NMR structure of a DNA kissing complex," *J. Biomol. Struct. Dyn.* **19**:649–658 (2002).

238. Hoogsteen, K., "The structure of crystals containing a hydrogen-bonded complex of 1-methylthymine and 9-methyladenine," *Acta Cryst.* **12**:822–823 (1959).
239. Roy, S., Papastavros, M. Z., and Redfield, A. G., "Procedure for C2 deuteration of nucleic acids and determination of A ψ 31 pseudouridine conformation by nuclear Overhauser effect in yeast tRNA^{Phe}," *Nucl. Acids Res.* **10**:8341–8349 (1982).
240. Abrescia, N. G. C., Gonz  les, C., Gouyette, C., and Subirana, J. A., "X-ray and NMR studies of the DNA oligomer d(ATATAT): Hoogsteen base pairing in duplex DNA," *Biochemistry* **43**:4092–4100 (2004).
241. Schmitz, U. and James, T. L. "Nuclear magnetic resonance and nucleic acids," in James, T. L., ed., *Methods in Enzymology*, Academic Press, San Diego, 1995, Vol. 261, p. 3–44.
242. Crippen, G. M. and Havel, T. F., *Distance Geometry and Molecular Conformation*, Research Studies Press, Taunton, MA, 1988.
243. Lane, A. N., Jenkins, T. C., Brown, T., and Neidle, S., "Interaction of berenil with the EcoRI dodecamer d(CGCGAATTCGCG)₂ in solution studied by NMR," *Biochemistry* **30**:1372–1385 (1991).
244. Hajduk, P. J., Meadows, R. P., and Fesik, S. W., "NMR-based screening in drug discovery," *Quart. Rev. Biophys.* **32**:211–240 (1999).
245. Meyer, B. and Peters, T., "NMR spectroscopy techniques for screening and identifying ligand binding to protein receptors," *Angew. Chem. Int. Ed.* **42**:864–890 (2003).
246. Meyer, B., Weimar, T., and Peters, T., "Screening mixtures for biological activity by NMR," *Eur. J. Biochem.* **246**:705–709 (1997).
247. Ni, F., "Recent developments in transferred NOE methods," *Prog. NMR Spectrosc.* **26**:517 (1994).
248. Tjandra, N. and Bax, A., "Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium," *Science* **278**:1111–1114 (1997).
249. Tjandra, N., Tate, S., Ono, A., Kainosho, M., and Bax, A., "The NMR structure of a DNA dodecamer in an aqueous dilute liquid crystalline phase," *J. Am. Chem. Soc.* **122**:6190–6200 (2000).
250. Wu, Z., Delaglio, F., Tjandra, N., Zhurkin, V. B., and Bax, A., "Overall structure and sugar dynamics of a DNA dodecamer from homo- and heteronuclear dipolar couplings and ³¹P chemical shift anisotropy," *J. Biomol. NMR* **26**:297–315 (2003).
251. Wertz, J. E. and Bolton, J. R., *Electron Spin Resonance*, McGraw-Hill, New York, 1972.
252. Gaffney, B. J. "Electron spin resonance of biomolecules," in Meyers, R. A., ed., *Encyclopedia of Molecular Cell Biology and Molecular Medicine*, Wiley-VCH, Weinheim, Germany, 2004, Vol. 4, pp. 115–133.
253. Biwas, R., K  hne, H., Brudvig, G. W., and Gopalan, V., "Use of EPR spectroscopy to study macromolecular structure and function," *Sci. Prog.* **84**:45–68 (2001).
254. Cantor, C. R. and Schimmel, P. R., *Biophysical Chemistry. Part II: Techniques for the Study of Biological Structure and Function*, Freeman, San Francisco, 1980.
255. Lipson, H. S., *Crystals and X-rays*, Wykeham Publications, London, 1970.
256. Rhodes, G., *Crystallography Made Crystal Clear*, Academic Press, San Diego, 2006.
257. Egli, M., "Nucleic acid crystallography: Current progress," *Curr. Opin. Chem. Biol.* **8**:580–591 (2004).

258. Peek, M. E., Williams, L. D. "X-ray crystallography of DNA-drug complexes," in Claies, J. B. and Waring, M. J., eds., "*Methods in Enzymology*, Academic Press, San Diego, 2001, Vol. 340, pp. 282–289.
259. Bragg, W. L., "The diffraction of short electromagnetic waves by a crystal," *Proc. Cambridge Phil. Soc.* **17**:43–57 (1914).
260. Ban, N., Nissen, P., Hansen, J., Moore, P. B., and Steitz, T. A., "The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution," *Science* **289**:905–920 (2000).
261. Nissen, P., Hansen, J., Ban, N., Moore, P. B., and Steitz, T. A., "The structural basis of ribosome activity in peptide bond synthesis," *Science* **289**:920–930 (2000).
262. Schuwirth, B. S., Borovinskaya, M. A., Hau, C. W., Zhang, W., Vila-Sanjurjo, A., Holton, J. M., and Cate, J. H. D., "Structures of the bacterial ribosome at 3.5 Å resolution," *Science* **310**:827 (2005).
263. Allen, M. P. and Tildesley, D. J., *Computer Simulation of Liquids*, Oxford Univ. Press, New York, 1987.
264. Frenkel, D. and Smit, B., *Understanding Molecular Simulations: From Algorithms to Applications*, Academic Press, San Diego, 1996.
265. McCammon, J. A. and Harvey, S. C., *Dynamics of Proteins and Nucleic Acids*, Cambridge Univ. Press, Cambridge, 1987.
266. Burkert, U. and Allinger, N. L., *Molecular Mechanics*, American Chemical Society, Washington, DC, 1980.
267. Rapaport, D. C., *The Art of Molecular Dynamics Simulation*, Cambridge Univ. Press, Cambridge, UK, 1995.
268. Torda, A. E., Scheek, R. M., and van Gunsteren, W. F., "Time-averaged nuclear Overhauser effect distance restraints applied to tendamistat," *J. Mol. Biol.* **214**:223–235 (1990).
269. Fennen, J., Torda, A. E., and van Gunsteren, W. F., "Structure refinement with molecular dynamics and a Boltzmann weighted ensemble," *J. Biomol. NMR* **6**:163–170 (1995).
270. Bonvin, A. M. J. J., and Brunger, A. T., "Conformational variability of solution nuclear magnetic resonance structures," *J. Mol. Biol.* **250**:80–93 (1995).
271. Kemmink, J. and Scheek, R. M., "Dynamic modelling of a helical peptide in solution using NMR data: Multiple conformations and multi-spin effects," *J. Biomol. NMR* **5**:33–40 (1995).
272. Pearlman, D. A. and Kollman, P. A., "Are time-averaged restraints necessary for NMR refinement? A model study for DNA," *J. Mol. Biol.* **220**:457–479 (1991).
273. Pearlman, D. A., "FINGAR: A new genetic algorithm-based method for fitting NMR data," *J. Biomol. NMR* **8**:49–66 (1996).
274. Gyi, J. I., Lane, A. N., Conn, G. L., and Brown, T., "Solution structures of DNA-RNA hybrids with purine-rich and pyrimidine-rich strands: Comparison with the homologous DNA and RNA duplexes," *Biochemistry* **37**:73–80 (1998).
275. Görler, A., Ulyanov, N. B., and James, T. L., "Determination of the populations and structures of multiple conformers in an ensemble from NMR data: Multiple-copy refinement of nucleic acid structures using floating weights," *J. Biomol. NMR* **16**:147–164 (2000).

276. This image was produced using the Chimera package from the Computer Graphics Laboratory, University of California, San Francisco: Huang, C. C., Couch, G. S., Petersen, E. F., and Ferrin, T. E., "Chimera: An extensible molecular modeling application constructed using standard components," *Proc. Pacific Symp. on Biocomputing*, 1996, Vol. 1, p. 724.
277. Berg, J. M., Tymoczko, J. L., and Stryer, L., *Biochemistry*, Freeman, New York, 2002.