

DNA: STRUCTURE AND FUNCTION

Richard R. Sinden, Christopher E. Pearson,
Vladimir N. Potaman, and David W. Ussery

I.	Introduction to the Structure, Properties, and Reactions of DNA	2
A.	Introduction	2
B.	The Structure of Nucleic Acids	3
C.	The Structure of Double-Stranded DNA	9
II.	DNA Curvature and Bending	16
A.	Introduction	16
B.	DNA Sequence Organization Required for Curvature	19
C.	Models for Curvature	20
D.	A Tract DNA Adopts a Unique Double Helical Conformation	22
E.	Which Model Best Explains DNA Curvature?	25
F.	Environmental Influences on DNA Curvature	26
G.	Proteins That Bind and Bend DNA	26
H.	The Biology of DNA Curvature	29
III.	Structure and Function of Supercoiled DNA	31
A.	Introduction	31
B.	Supercoiled Forms of DNA	31
C.	The Biology of Supercoiled DNA	38
IV.	Cruciform Structures	42
A.	Introduction	42
B.	Formation and Stability of Cruciforms	43

Advances in Genome Biology

Volume 5A, pages 1-141.

Copyright © 1998 by JAI Press Inc.

All rights of reproduction in any form reserved.

ISBN: 0-7623-0079-5

C.	Cruciform Structure	46
D.	Assays for Cruciform Structures in DNA.....	50
E.	The Biology of Inverted Repeats	53
V.	Left-Handed Z-DNA.....	64
A.	Introduction.....	64
B.	The Structure of Z-DNA.....	65
C.	Formation and Stability of Z-DNA.....	69
D.	Assays for Z-DNA	71
E.	Z-DNA <i>In Vivo</i>	72
F.	Possible Biological Functions of Z-DNA.....	73
VI.	Triple-Stranded Nucleic Acid.....	75
A.	Introduction.....	75
B.	The Structure of Triplex DNA	75
C.	The Formation and Stability of Triplex DNA.....	80
D.	Assays for Triplex DNA	84
E.	Triplex DNA <i>In Vivo</i>	85
F.	Possible Biological Roles of H-(H*)-DNA	88
G.	Control of Gene Expression	96
VII.	Miscellaneous Alternative Conformations of DNA	105
A.	Slipped Strand DNA.....	105
B.	DNA Unwinding Elements.....	107
C.	Parallel-Stranded DNA	108
D.	Four-Stranded DNA	109
E.	Higher Order Pu-Py Structures	113

I. INTRODUCTION TO THE STRUCTURE, PROPERTIES, AND REACTIONS OF DNA

A. Introduction

DNA occupies a critical role in the cell, inasmuch as it is the source of all intrinsic genetic information. Chemically, DNA is a very stable molecule, a characteristic important for a macromolecule that may have to persist in an intact form for a long period of time before its information is accessed by the cell. Although DNA plays a critical role as an informational storage molecule, it is by no means as unexciting as a computer tape or disk drive. Rather, DNA can adopt a myriad of alternative conformations, including cruciforms, intramolecular triplexes, left handed Z-DNA, and quadruplex DNA, to name a few. Local variations in the shape of the canonical B-form DNA helix are most certainly important in DNA-protein interactions that modulate and control gene expression. Moreover, the ability of DNA to adopt many alternative helical structures, the ability to bend and twist, and the ability to modulate the potential energy of the molecule through variations in DNA supercoiling provide enormous potential for the involvement of

the DNA itself in its own expression and replication. This chapter will focus on alternative structures of DNA and their potential involvement in biology. For more detail on some subjects, see books by Sinden¹ and Soyfer and Potaman.²

B. The Structure of Nucleic Acids

3. Bases

Two different heterocyclic aromatic bases with purine heterocycles, adenine and guanine, exist in DNA (Figure 1). Adenine has an amino group (-NH₂) at the C6 position, whereas guanine has an amino group at the C2 position and a carbonyl group at the C6 position. Two pyrimidine bases, thymine and cytosine, are commonly found in DNA. Thymine contains a methyl group at the C5 position, with carbonyl groups at the C4 and C2 positions. Cytosine contains a hydrogen atom at the C5 position, with an amino group at C4. Uracil, which is used in place of thymine in RNA, lacks the methyl group at the C5 position. Uracil is not usually found in DNA, but can result from cytosine deamination. The purines and pyrimidines are excellent candidates for informational molecules. The specific placement of hydrogen bond donor and acceptor groups provides unique structural identity. The hydrogen atoms of amino groups provide hydrogen bond donors, and the carbonyl oxygen and ring nitrogens provide hydrogen bond acceptors.

2. Deoxyribose Sugar

β -D-2-Deoxyribose is a flexible and dynamic part of the DNA molecule (Figure 2A). A shift in the positions of the C2' and C3' carbons relative to a flat plane through all carbon atoms results in various twist forms of the sugar ring. Several sugar conformations are found in DNA, the most common of which are the C2' endo and C3' endo forms (Figure 2B).

3. Nucleosides and Nucleotides

Nucleosides (adenosine, guanosine, thymidine, and cytidine) are composed of a base and a deoxyribose sugar. Nucleotide refers to the base, sugar, and phosphate group. The phosphate group is attached to the 5' carbon of the deoxyribose (Figure 3). One, two, or three phosphate groups on a sugar are designated as α , β , and γ , for the first, second, and third, respectively (Figure 3). A phosphate group can also be attached to the 3' or 5' carbon of deoxyribose.

The glycosidic bond is the bond between the sugar and the base. In the α configuration, the bond is on the 3'-OH side of the ribose sugar. This is in contrast to the β , where it is on the 5'-OH side. The base can rotate around the glycosidic bond, but generally it exists in one of two standard conformations: *syn* and *anti*. The *anti* conformation reflects the relative spatial orientation of the base and sugar

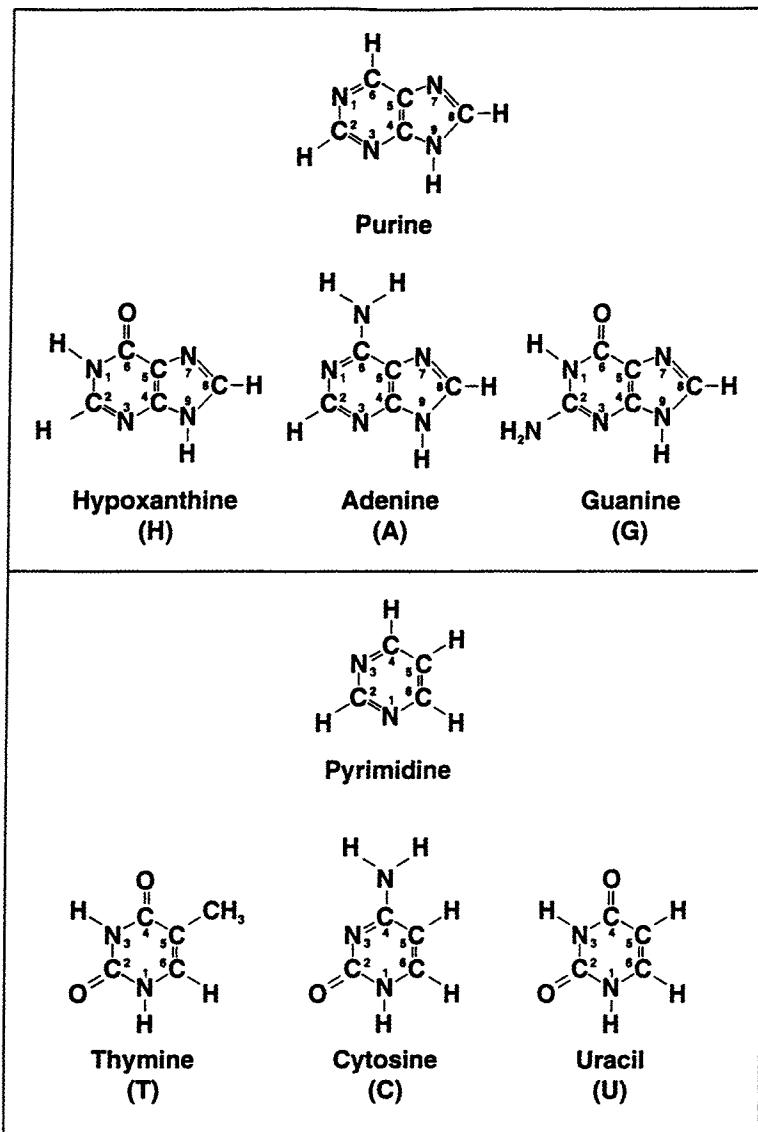


Figure 1. Purine and pyrimidine bases. **(Top)** The two-member purine aromatic ring consists of fused six- and five- member rings, each composed of carbon and nitrogen. The structures and position numbers for the basic purine ring, hypoxanthine (H), adenine (A), and guanine (G), are shown. **(Bottom)** The aromatic pyrimidine ring is composed of six carbon and nitrogen atoms. The basic ring structures, thymine (T), cytosine (C), and uracil (U), are shown.

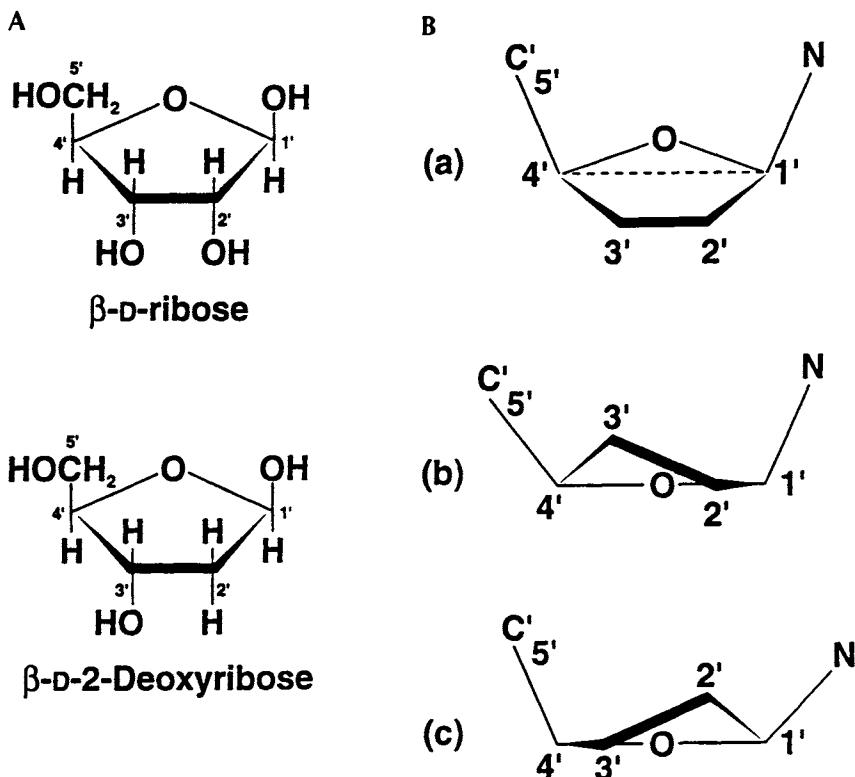


Figure 2. Sugars associated with DNA. (A) β -D-Ribose is found in RNA molecules. β -D-2-Deoxyribose, found in DNA, lacks a hydroxyl at the 2' position. The positions of the carbon atoms in the ribose ring are numbered with primes (e.g., 2'). (B) The sugar residue can adopt many different twist conformations. (a) Representation of an envelope conformation of a ribose sugar. (b) Representation of the C3' endo conformation of the ribose sugar. (c) Representation of the C2' endo ribose sugar conformation.

as found in most conformations of DNA, in which the ring is away from the ribose. The *syn* conformation, in which the ring is spatially over the ribose, is found in the Z-form DNA.

4. The Phosphodiester Bond

In DNA (and in RNA) nucleotides are joined by a 3'- 5' phosphodiester bond that connects the 3' sugar carbon of one nucleotide to the 5' sugar carbon of the adjacent nucleotide through the phosphate (Figure 4). (This is termed the 3'-5' phosphodiester bond.) At the physiologically important pH 7, the ionized phos-

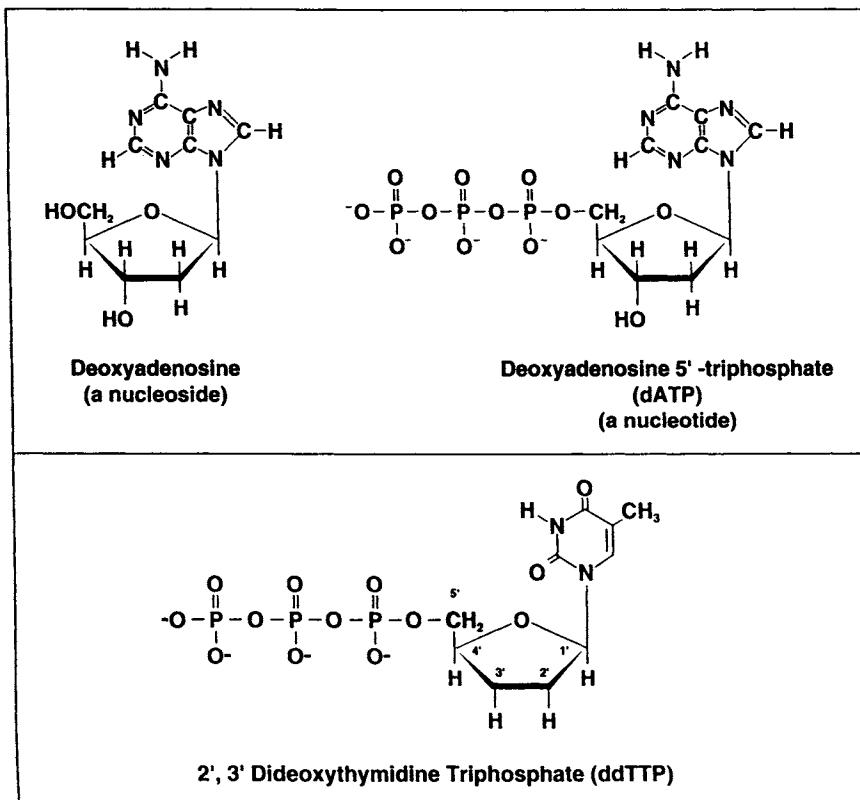


Figure 3. Nucleosides and nucleotides. The nucleoside deoxyadenosine consists of adenine linked through the C—N glycosidic bond to the C1' position of a 2' deoxyribose sugar. The nucleotide deoxyadenosine 5'-triphosphate (dATP) consists of adenine linked to a deoxyribose 5'-triphosphate. 2', 3'-Dideoxythymidine triphosphate (ddTTP) contains no hydroxyl group on either the 2' or 3' positions. Dideoxynucleotides are used for DNA sequencing reactions, since DNA polymerases require a 3'-OH for the addition of the next deoxyribonucleotide.

phate groups have one negative charge per nucleotide, which creates repulsive forces between complementary polynucleotide strands.

An important point regarding the structure of a polynucleotide is that it has two distinct ends called the 5' and 3' ends. These different ends define a polarity to the individual strands of DNA. Frequently, a hydroxyl group exists at 3' ends (3'-OH) and a single phosphate group at 5' ends (5'-PO₄). DNA replication and transcription occur by the addition of nucleoside 5' triphosphates to the 3' hydroxyl group of the terminal nucleotide of the polynucleotide.

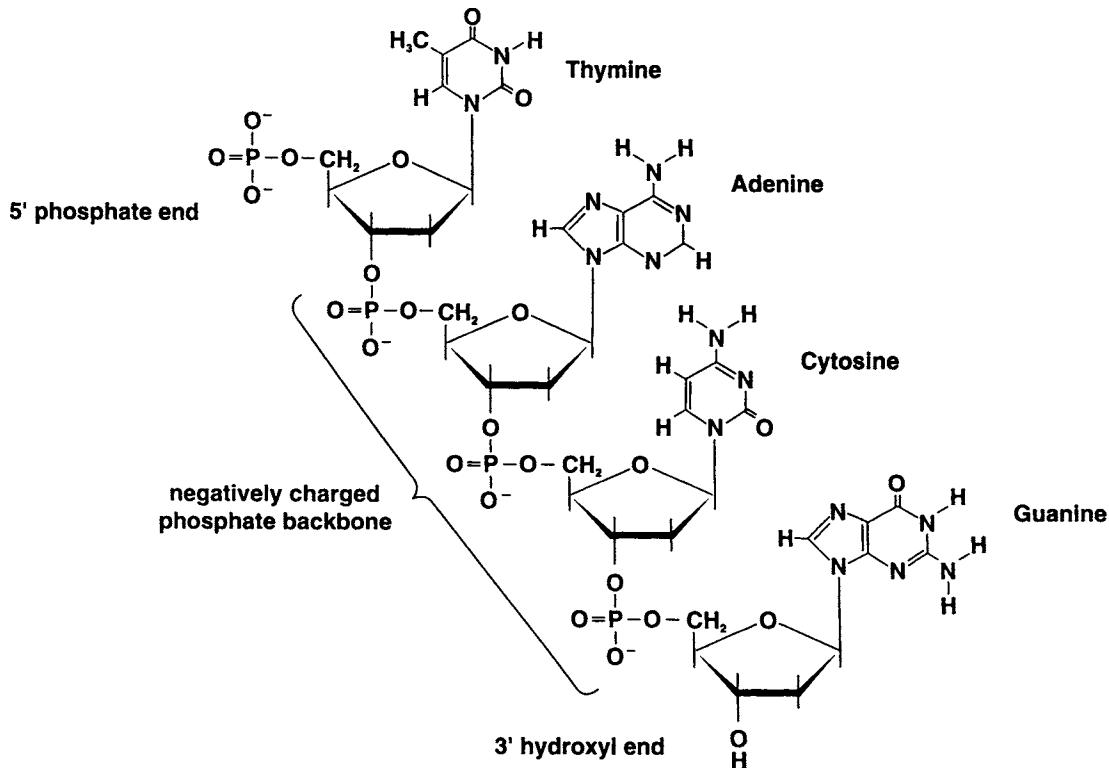
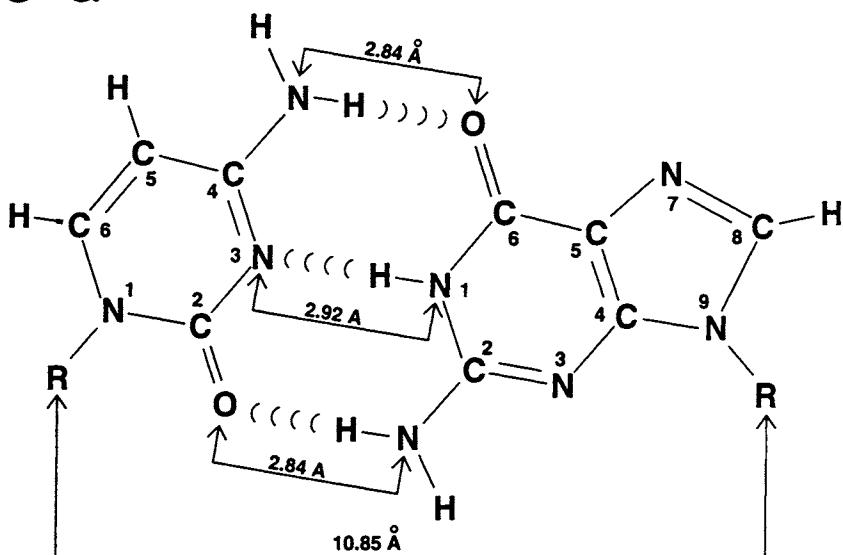
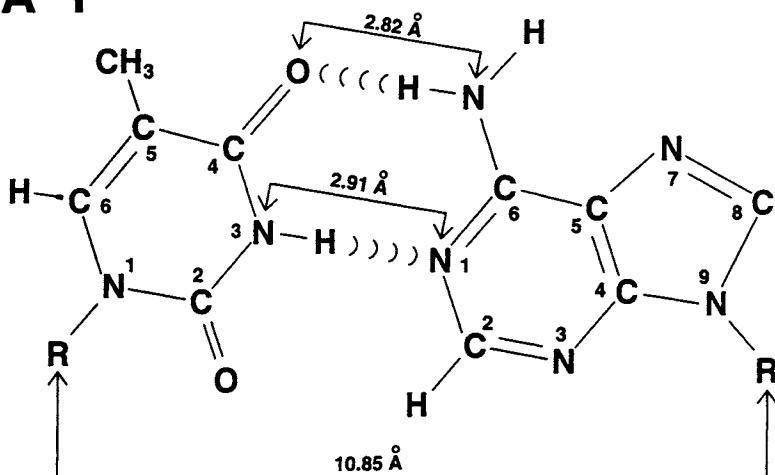


Figure 4. The phosphodiester bond. A single strand of DNA consists of nucleosides linked through a phosphate covalently bound to their 5' and 3' carbons. The charged phosphate groups are responsible for the negative charge on DNA. The two ends of the DNA chain are chemically distinct, typically having a 5'- PO_4 and 3'-OH.

A

C·G**A·T**

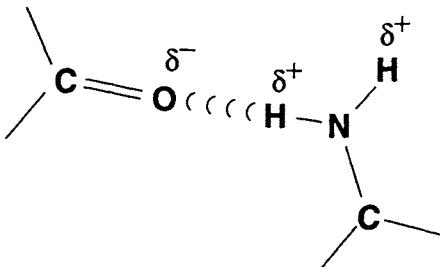
B

Figure 5. Hydrogen bonds in Watson–Crick base pairs. (A) The interatom distances between the C1' position of the ribose sugars are indicated. The distance between nitrogens or nitrogen and oxygen involved in hydrogen bonding are also shown. The curves in the direction of the hydrogen bond acceptor (N or O atoms) represent the hydrogen bonds. (B) Schematic of a hydrogen bond in which two electronegative atoms (N and O) partly share a proton of the NH₂ group. In this scheme the nitrogen atom is a donor and the oxygen atom is an acceptor of the hydrogen bond.

C. The Structure of Double-Stranded DNA

The structure of the DNA described by Watson and Crick in 1953 is a right handed helix of two individual *antiparallel* DNA strands. Hydrogen bonds provide specificity that allows pairing between the complementary bases (A-T and G-C) in opposite strands. Base stacking occurs near the center of the DNA helix and provides a great deal of stability to the helix (in addition to hydrogen bonding). The sugar and phosphate groups form a “backbone” on the outside of the helix. There are about 10 base pairs (bp) per turn of the double helix.

1. Hydrogen Bonding and Base Stacking

Hydrogen bonds are short, noncovalent, directional interactions of 2.6–3.1 Å between a H atom (a donor), with a partial positive charge and a negatively charged acceptor atom, usually a carbonyl oxygen (-C=O) or nitrogen (N:) (Figure 5). In the DNA double helix, the N and O atoms involved in hydrogen bonding are separated by about 2.8–2.92 Å. An A-T base pair has two hydrogen bonds separated by 2.82 and 2.91 Å, whereas a G-C base pair has three hydrogen bonds separated by 2.84–2.92 Å (Figure 5).^{3,4} Hydrogen bonds have an energy of about 3–7 kcal/mol. By contrast, covalent bonds are equivalent to 80–100 kcal/mol. Hydrogen bonds can be deformed by stretching and bending. The energy of hydrogen bonds in DNA is about 2–3 kcal/mol which is weaker than most hydrogen bonds. This is due to geometric limitations within the double helix that preclude an optimal directional alignment of the bonds.

Hydrophobic and Van der Waals interactions are involved in stacking between the aromatic planar bases. Stacking interactions are estimated to be about 4–15

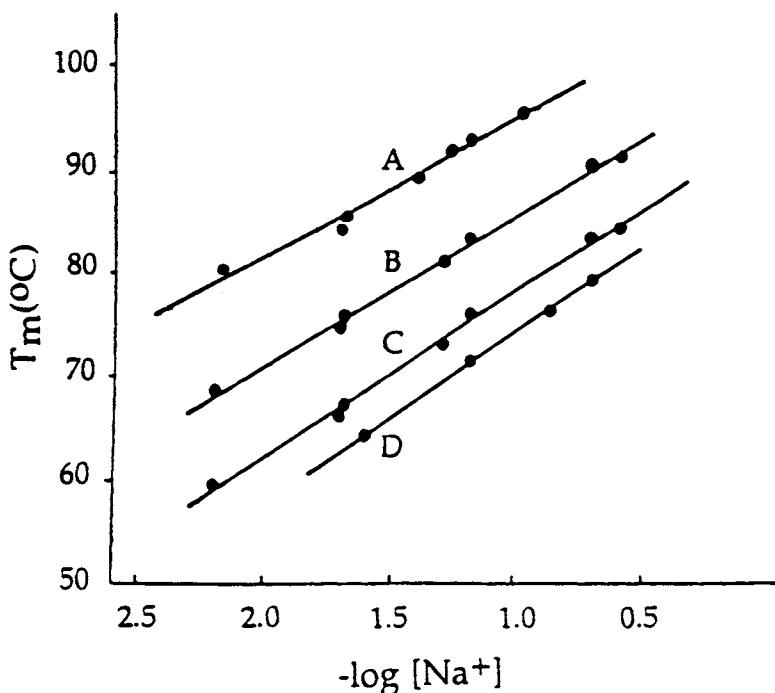


Figure 6. Dependence of DNA melting temperature (T_m) on the concentration of sodium ions in the medium. In accordance with the higher number of hydrogen bonds in the G-C pair as compared with the A-T pair, the four lines also show the dependence of T_m on the percentage of G-C pairs. (A) %G-C = 72; (B) %G-C = 50; (C) %G-C = 33; (D) %G-C = 24. Adapted from Frank-Kamenetskii.⁵

kcal/mol per dinucleotide. Base stacking provides energies of stabilization similar to those provided by hydrogen bonding.

Differences between the base stacking and hydrogen bonding energies for individual dinucleotides contribute to the heterogeneity of the B-form DNA helix. The overall energy of hydrogen bonding depends predominantly on base composition. That is, individual hydrogen bonds in A-T or G-T base pairs have relatively the same geometry and strength of hydrogen bonding. On the other hand, base stacking energies depend on the sequence of the DNA. For example, a 5' GT and 5' TG dinucleotides have very different stacking energies of 10.51 kcal/mol and 6.78 kcal/mol, respectively (see Sinden,¹ Table 1.2).

Hydrogen bonding and base stacking contribute to the stability of the DNA double helix. Since the energies for both stacking and hydrogen bonding are greater for G+C-rich DNA, it may not be surprising that the melting temperature (T_m) of the DNA is a function of G + C content of the DNA (Figure 6).⁵

2. Non-Watson-Crick Bonds

There are many ways in which two bases can form hydrogen bonds. Several of these are shown in Figure 7. Reverse Watson–Crick base pairs have one nucleotide rotated 180° with respect to the complementary nucleotide, relative to the Watson–Crick structure. In this structure, the glycosidic bonds are in a *trans* rather than *cis* orientation. Because of symmetrical hydrogen bonding potential at the C2-N3-C4 positions, thymine can rotate at the N3-C6 axis to form a reverse Watson-Crick A·T base pair. Hoogsteen base pairs utilize the C6-N7 face of the purine for hydrogen bonding with the N3-C4 face of the pyrimidine.⁶ In Hoogsteen base pairing, the N7 position of purine is base-paired, altering the chemical reactivity of this position, relative to that expected for a Watson–Crick base pair. In a reverse Hoogsteen base pair, one of the bases is positioned 180° with respect to the other base compared to the normal Hoogsteen bond. A number of other base pairing schemes are possible.

3. Keto-enol Tautomerizations Can Result in Non-Watson-Crick Base Pairs

The C6 keto (C=O) position of guanine and the C4 keto of thymine can undergo a tautomerization to an enol form (C-OH). For this tautomerization to occur, the double bond must shift from the carbonyl group to the nitrogen–carbon bond in the ring. In a similar fashion, an amino nitrogen (-NH₂) can undergo a transition to an imino form (=NH). This can occur at the C6 position of adenine or the C4 position of cytosine. The imino or enol forms of the bases each have two isomeric forms that can exist.

The chemical equilibrium between the alternative tautomeric forms favors the keto and amino forms by about 10⁴. Tautomerization leads to a reversal of the polarity of hydrogen bonding, which can result in mispairing. Enol-G can pair with T, keto-T can pair with G, imino-A can pair with C, and imino-C can pair with A. There is little evidence so far to suggest that keto-enol and amino-imino tautomerizations occur during the synthesis of DNA *in vivo*. Mispairs *in vivo* may arise from ionized and wobble base pairs.

When bases become ionized, their hydrogen bonding properties are changed. This can lead to many non-Watson–Crick base pairs. Adenine is prone to protonation at low pH, which can lead to the formation of an A⁺·C wobble base pair. Cytosine is also very prone to protonation which can lead to a C⁺·G Hoogsteen base pair. The ionized form of thymidine can form a T·G base pair.

4. B-form DNA

The structure of B-form DNA was determined from X-ray diffraction analysis of the sodium salt of DNA fibers at 92% relative humidity.^{7,8} B-form DNA is shown schematically in Figure 8. There are about 10.0 bp per right-handed helical

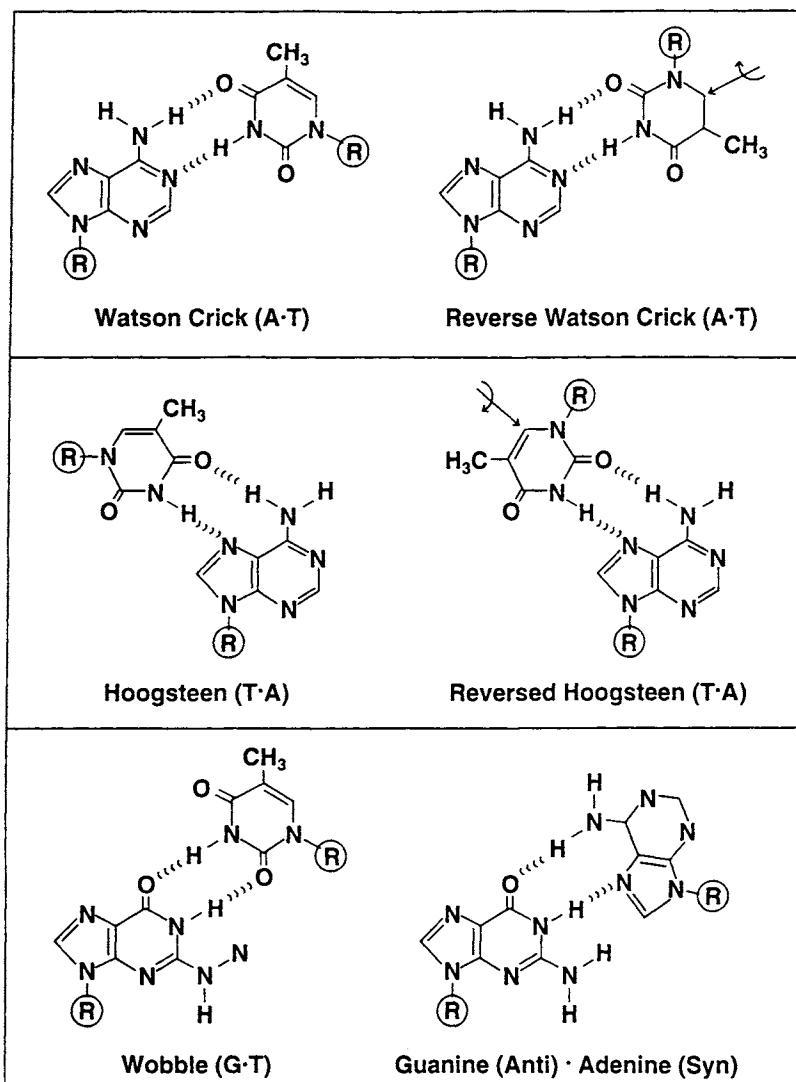


Figure 7. Base-pairing schemes. The A-T Watson-Crick base pair (top left) differs from a reversed Watson-Crick pair by the 180° rotation of the pyrimidine base (top right). T and A can also form Hoogsteen and reversed Hoogsteen base pairs (middle), with hydrogen bond formation between the pyrimidine and the N1, C6, N7 face of the purine base. A 180° rotation of the pyrimidine is required for the formation of the reversed Hoogsteen base pair. In the wobble G-T base pair (bottom left) the pyrimidine is shifted up vertically. The G(anti)·A(syn) base pair (bottom right) involves pairing between two purines, using the Watson-Crick surface of G (in the typical *anti* conformation) and the Hoogsteen surface of A (in the *syn* conformation).

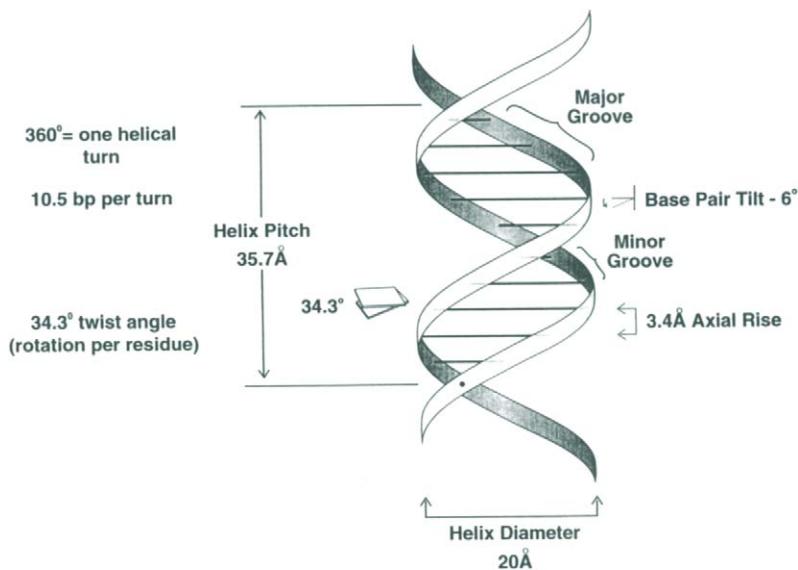


Figure 8. Structural parameters of B-form DNA. B-form DNA contains about 10.5 bp/helical turn, a 34.3° twist angle, a helix pitch of 35.7 \AA , an axial rise of 3.4 \AA , a tilt angle of about -6° , and a helix diameter of about 20 \AA . The major and minor grooves are indicated.

turn in B-DNA in fibers. Helix parameters are defined in Table 1 and Figure 9 and listed in Table 2. The form of the ribose sugar is C2' endo (Figure 2). The term “B-form DNA” will be used to refer to the right-handed helical form commonly found for DNA in solution, where the helix repeat is 10.5 bp.

A major feature of B-form DNA is the presence of two distinct grooves, a major and a minor groove, shown in Figure 8. These two grooves provide very well defined surfaces with different shapes and geometries of hydrogen bonding potentials. The grooves are used for differential protein interaction. Certain DNA binding proteins and chemicals interact with either the major or minor groove. The Watson–Crick hydrogen bonding surfaces are not available to solvent or proteins, since they participate in hydrogen bonding with each other at the center of the double helix. The Hoogsteen hydrogen bonding surface of purines is accessible through the major groove in B-form DNA.

5. A-form DNA

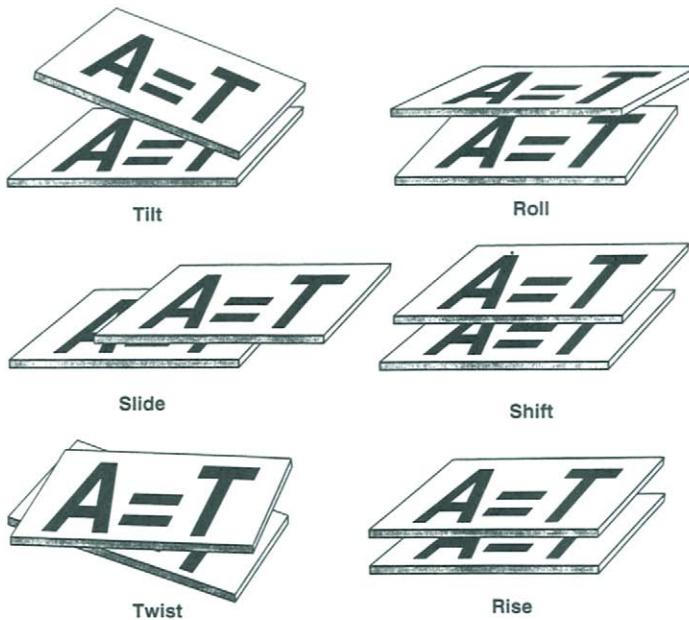
A-form DNA was originally identified using X-ray diffraction in DNA fibers at 75% relative humidity.⁹ The grooves in A-DNA are not as deep as those in B-DNA,

Table 1. Definitions of Helical Characteristics

<i>Characteristic</i>	<i>Definition</i>
Helix sense	The helical rotation of the double helix. The structure described by Watson and Crick is a right-handed (clockwise) helix. Most helical forms of DNA are right-handed. Left-handed DNA, called Z-DNA, is discussed below.
Residues per turn	The number of base pairs in one helical turn of DNA, i.e., the number of bases needed to complete one 360° rotation. DNA in solution contains about 10.4–10.5 bp/turn, although this value can vary considerably as a function of base composition.
Axial rise	The distance between adjacent planar bases in the DNA double helix. In B-form DNA there are about 3.4 Å between adjacent base pairs.
Helix pitch	The length of one complete helical turn of DNA. In B-form DNA in solution, one helical turn of 10.5 bp is completed in 36 Å.
Base pair tilt	The angle of the planar bases with respect to the helical axis. The tilt angle is measured by considering the angle made by a line drawn through the two hydrogen-bonded bases relative to a line drawn perpendicular to the helix axis. For a base pair perpendicular to the helix axis, the tilt angle is 0°. In B-form DNA the bases are tilted by -6°. In A-form DNA the base pairs are significantly tilted at an angle of 20°.
Base pair roll	The angle of deflection of a base pair with respect to the helix axis along a line drawn between two adjacent base pairs relative to a line drawn perpendicular to the helix axis. Compare this to the tilt in which the line angle is measured along a line drawn through the base pair. The roll and tilt angles are offset by 90°.
Propeller twist	The angle between the planes of two paired bases. A base pair is rarely a perfect flat plane with each aromatic base in the same plane. Rather, each base has a slightly different roll angle with respect to the other base. This makes the two bases look like an airplane propeller.
Helix diameter	Diameter of the helix refers to the width in Å across the helix. B-DNA has a diameter of 20 Å.
Rotation per residue (twist angle)	The angle between two adjacent base pairs. Consider the angle between lines drawn through two adjacent base pairs. In B-form with 10 bp in one 360° helix turn of DNA, the rotation per residue is 36°. For B-form in solution with 10.5 bp/turn, $h = 34.3^\circ$.
Shift	The displacement of two bases in the direction of the major or minor groove.
Slide	The displacement of two bases in the direction of the phosphate backbone.

and the bases are tilted to about 20°. Moreover, in A-DNA sugar pucker is C3' endo compared to C2' endo for B-DNA. Runs of homopurine/homopyrimidine DNA sequence [poly (dG)-poly (dC), for example] seem to set up an A-like helix, as determined by characteristic circular dichroism (CD) spectra.¹⁰ Therefore, within a B-like DNA molecule, specific regions may exist in an A-DNA form. Many regions

Spatial Relationship Between Adjacent Base Pairs



Spatial Relationship Between Bases in a Base Pair

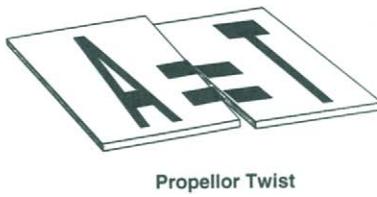


Figure 9. Spatial relationships between adjacent base pairs. Representations of the parameters described in Table 1.

of RNA molecules, including transfer RNA (tRNA), ribosomal RNAs (rRNA), and in parts of messenger RNAs (mRNA), exist in a double-helical A-form.

6. Sequence-Dependent Variation in the Shape of the DNA

A canonical textbook B-form helix is not likely to exist in nature. The actual shape of DNA will depend on its base composition, the local flanking sequence,

Table 2. Helix Parameters

Parameter	A-DNA	B-DNA	Z-DNA
Helix sense	Right	Right	Left
Residue per turn	11	10.5	12
Axial rise	2.55 Å	3.4 Å	3.7 Å
Helix pitch	28 Å	36 Å	45 Å
Base pair tilt	20°	-6°	7°
Rotation per residue	33°	34.3°	-30°
Diameter of helix	23 Å	20 Å	18 Å
Glycosidic bond configuration			
dA, dT, dC	<i>anti</i>	<i>anti</i>	<i>anti</i>
dG	<i>anti</i>	<i>anti</i>	<i>syn</i>
Sugar pucker			
dA, dT, dC	C3' endo	C2' endo	C2' endo
dG	C3' endo	C2' endo	C3' endo
Intrastrand phosphate-phosphate distance			
dA, dT, dC	5.9 Å	7.0 Å	7.0 Å
dG	5.9 Å	7.0 Å	5.9 Å

and environmental conditions. Evidence for local structural variation in DNA comes from X-ray crystallography.^{11,12} The twist angle between adjacent base pairs can vary considerably, from 32° to 45°.¹¹ Because of flanking sequence effects, not all dinucleotides will have the same twist angle. Since individual twist angles are different, the actual helical repeat and therefore the exact shape of a 10.5-bp helical turn of DNA can vary considerably.

Polymeric regions of a single base in one strand can adopt unusual helical forms. Poly(dA)-poly(dT) forms an unusual structure called heteronomous DNA,¹³ with a helix repeat of 10 bp/turn. In heteronomous DNA the deoxyribose sugar in the d(A) strand is C3' endo, whereas the deoxyribose in the d(T) strand is C2' endo. The helical changes associated with phased runs of A in DNA, which are responsible for DNA bending, are another example of a variation in the DNA structure. A-form-like tracts of DNA form in runs of poly(dG)-poly(dC). Runs of (dG)-(dC) greater than 20 bp can form triple-stranded and four-stranded structures, as discussed below. Triplet repeats can form an unusual helix structure that may be more flexible than canonical B-form DNA.^{14,14a}

II. DNA CURVATURE AND BENDING

A. Introduction

In some respects DNA is a simple molecule. A plasmid DNA of known sequence can be cut into a number of different sized pieces by a restriction enzyme. Molecules of equal length would be expected to exhibit similar flexibil-

ity. Pieces of DNA shorter than the persistence length of 150–200 bp will behave as rather stiff rods that cannot be easily bent into a circle (see refs 15 and 16 for a discussion of persistence length). Larger pieces adopt a “random coil” shape in solution. DNA molecules of defined size behave in a very predictable way when run on agarose or acrylamide gels. Shorter molecules migrate faster than larger molecules. In both agarose and acrylamide gels there is a linear relationship between the log of the distance migrated and the length of the DNA in base pairs.

Certain DNA fragments of known length do not run at their expected position on an acrylamide gel. One of the most striking examples of such anomalous migration was a 414-bp piece of kinetoplast DNA from *Crithidia fasciculata*.^{17,18} This DNA migrated as if it were twice as long (e.g., it ran as if it were about 830 bp) in an acrylamide gel, but migrated at its proper position (414 bp) in an agarose gel. The anomalous migration of the kinetoplast DNA in polyacrylamide gels was attributed to the kinetoplast DNA being either stably curved or kinked (Figure 10). Although the migration through acrylamide gels is not completely understood in physical terms, the migration is believed to depend on the ability of DNA to “snake” through the gel matrix.¹⁹ A relatively straight piece of DNA that is rather flexible can easily snake through the gel. On the other hand, DNA containing a permanent bend or kink is not as flexible and can get hung up in the gel matrix. For DNA of equal sizes, a curved fragment will take longer to snake through the small pores or matrix of the acrylamide gel than a noncurved DNA fragment. (This does not occur in agarose gels because the matrix or pore size is believed to be larger.)

The term “curved DNA” will be used to describe the phenomenon of intrinsic DNA curvature, and the term “bent DNA” will refer to axis deflection introduced at one site in DNA (frequently introduced by some external agent such as thermal motion, proteins, drugs, etc.). The investigation of behavior of DNA fragments in polyacrylamide gels has become a principal instrument in the studies of curved DNA, although this is by no means the only method used to measure DNA curvature. Indeed, as will be seen below, various techniques give different results, which in turn are interpreted in terms of very different models. However, one important observation that all techniques (including, perhaps surprisingly to some, gel electrophoresis) agree on is that nearly any DNA sequence is capable of exhibiting some curvature, under the appropriate conditions, and that in fact it is quite rare to find a piece of DNA that is not curved to some extent. Most DNA found in organisms has a gentle left-handed writhe, due to slight intrinsic curvature^{20,21} (see Travers²² for a review). Perhaps this is not surprising, when one considers that the DNA in all organisms, from bacteria to humans, must be condensed by about 1000-fold to fit inside the cell. For bacteria, this means that a typical bacterial operon, if it were not curved, but stretched out in its B-DNA conformation, would be longer than the bacterial cell! Thus DNA curvature plays an important role in controlling the condensation as well as the expression of genetic information.

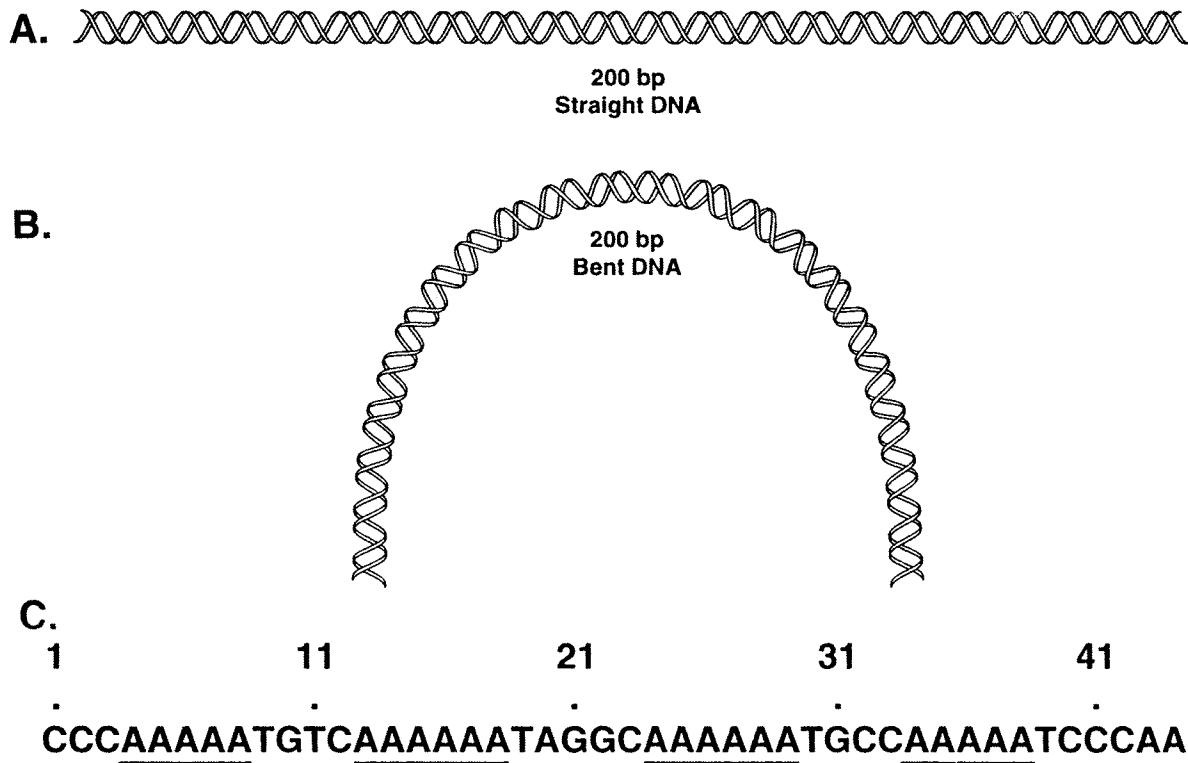


Figure 10. Curved DNA. (A) A 200-bp straight DNA molecule. (B) A 200-bp piece of curved DNA. (C) Phased A tracts are preset at the site of bending in a *Critchidia fasciculata* DNA sequence.

B. DNA Sequence Organization Required for Curvature

Wu and Crothers devised a clever way to map the site of curvature in the kinetoplast DNA.²³ They cloned the bent kinetoplast DNA fragment as a dimer of a 241-bp *Hind*III DNA fragment. By cutting the dimer with a series of different restriction enzymes, DNA fragments with the curve located at various positions within the 214-bp fragment were produced. The sequence organization (i.e., the location of the curve) of each fragment will be different from that of the original fragment. Such a set of DNA molecules is called *circularly permuted*.

Circularly permuted molecules do not migrate at the same rate because the center of curvature is at different locations within the molecule. This creates different end-to-end distances. The end-to-end distance is shortest when the curve is positioned at the center of the molecule. For molecules of the same size, the end-to-end distance is important in determining gel mobility. The molecule with the shortest end-to-end distance will migrate most slowly in an acrylamide gel. This pattern was observed by Wu and Crothers.²³

The DNA sequence shown in Figure 10B was found at the center of the bend. There are five runs of four or five A's, which, in each case, are preceded by a C and followed by a T. In addition, the runs of A are phased by 10 bp. The DNA curvature hypothesis suggested that the runs of A and the 10-bp phasing were important in curvature.^{23,24} Crothers also suggested that bending may occur at a junction resulting from the interruption of B-form DNA by an A tract that adopts a non-B-DNA helix.

The 10-bp phasing is required for curvature associated with A tracts. For DNA to contain a region of stable curvature and produce the anomalous gel mobility, the small individual curves associated with A tracts must be oriented in the same direction or same plane. If the A-tracts are placed every 1.5 turns of the DNA helix (15–16 bp), the curvature resulting from the individual A tracts will have a zigzag shape, and the DNA will essentially migrate as nonbent DNA on an acrylamide gel.

The phasing hypothesis was tested by Hagerman using oligonucleotides of sequence G_{A3}T₃C, G₂A₃T₃C₂, and G₃A₃T₃C₃, which phases the A₃ block in one DNA strand by 8-, 10-, and 12-bp intervals.²⁵ Only the G₂A₃T₃C₂ polymer with A tracts phased at 10 bp exhibited the pattern of electrophoretic migration diagnostic for curved DNA.

Koo, Wu, and Crothers²⁴ also synthesized a large number of oligonucleotides containing various lengths of A tracts that were phased at different lengths. Polymers with A_{4–9} were bent, with curvature being optimal for A₆. A continuous run of A's is required for curvature, since interruption of an A₅ tract with C, G, or T destroys the curvature. There is no particular sequence requirement for the base 5' or 3' to an A tract for curvature, although flanking sequences can influence curvature.

DNA sequences that do not contain runs of A's can also be curved. The curvature observed in DNA lacking phased A tracts is usually not as large as A tract curvature (in the absence of divalent cations). However, based on gel mobility experi-

ments Brukner et al. have found that, in the presence of Mg^{2+} or Ca^{2+} , certain sequences without A tracts (e.g., $N_4G_3C_3$) can exhibit strong macroscopic curvature.^{26,27} Furthermore, Brukner et al. have postulated that the direction of curvature was opposite that of the A tract curves.

C. Models for Curvature

1. *The Wedge Model for DNA Curvature*

Trifonov proposed a model for curvature, in which a wedge angle is associated with the AA dinucleotide and curvature is attributed to the summation of the wedge angles of the AA dinucleotides.^{28,29} The sum of wedges pointing in the same direction, a condition met by the 10-bp phasing, leads to the curvature of DNA. Ulanovsky et al. used measurement of the efficiency of ligation of small DNA molecules into circles to calculate the wedge angle of an AA dinucleotide.²⁹ As a short piece of DNA with a defined curvature is ligated together into increasingly long polymers, at some length the total angle of curvature will result in the formation of a circle of DNA. By determining the length at which the DNA forms a circle and by knowing the number of AA dinucleotides responsible for the 360° curvature, the individual AA wedge angle can be determined, assuming no contributions from the other DNA sequences present. The optimal size for circle formation in a polymer studied by Ulanovsky et al. was 126 bp. This is much less than the dynamic persistence length of DNA of about 230 bp estimated by electron microscopy.¹⁶ For a discussion of static versus dynamic persistence length and their relationship to DNA curvature, see Trifonov et al.³⁰ Thus efficient circularization is an indication that this sequence can readily adopt a structure in which it is curved by 360°. Within the 126 bp molecule there are 66 AA dinucleotides, and the sum of their individual wedge angles was assumed to be responsible for the 360° curvature. Trifonov estimated a total wedge angle (of both tilt and roll components) of 8.7° for each AA dinucleotide. This angle probably represents an upper limit of the wedge angle, and a later study estimated the AA dinucleotide wedge to be close to 1.1°,²⁰ in the absence of Mg^{2+} (Mg^{2+} is necessary for the ligation experiments) a value of roughly twice this was obtained for the same sequence, based on ligation experiments.³¹

2. *The Junction Model for DNA Curvature*

Wu and Crothers²³ proposed the junction model for DNA curvature (Figure 11), which suggests that there is a bend at the junction of B-form DNA and a non-B DNA helix associated with A tracts. A tracts can adopt a non-B-DNA helix called heteronomous DNA.¹³ In addition, modeling studies suggested that DNA would

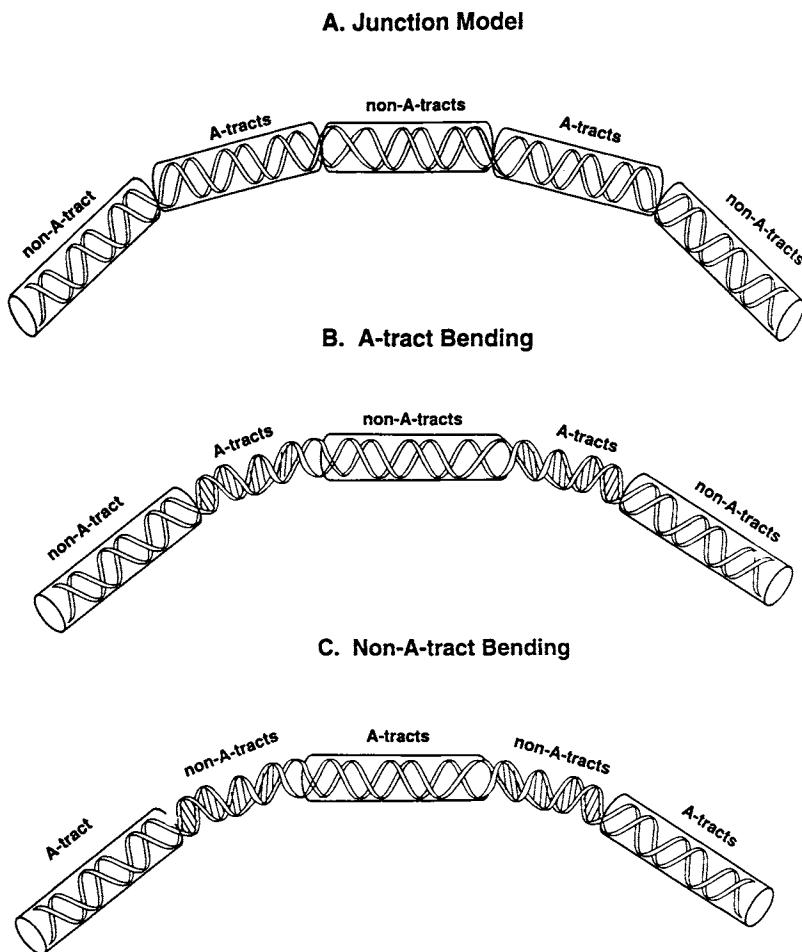


Figure 11. The wedge model and junction models for DNA bending. (A) The wedge model suggests that curvature results from the sum of the wedge angles present between adjacent AA dinucleotides within an A tract. (B) The junction model suggests that bending occurs at the junction between an A tract (which has a non-B-DNA helix) and mixed-sequence DNA that forms a B-DNA helix. (C) The general sequence model suggests that bending results from curvature present in mixed-sequence B-form DNA interspersed with A tracts that, in this model, are considered to be straight.

bend at the junction between A-DNA and B-DNA.³² There are two junctions, a 5' and a 3' junction, associated with each A-tract, and the site of curvature could be associated with either or both junctions. Data suggested that curvature was probably localized at the 3' end of the A tract.^{23,24}

3. *The General Sequence (or Non-A Tract) Model of DNA Curvature*

With the exception mentioned above on non-A tract DNA curvature, it might seem pretty conclusive that the A tracts have been shown to be clearly responsible for most DNA curvature observed in the gel electrophoresis. Indeed, it is possible to predict the anomalous mobility of a particular DNA fragment on an acrylamide gel, based on the sequence, using any one of several algorithms. Most of these models assume that the curvature is due to phased A tracts, as explained above. Thus, when Goodsell et al.³³ proposed a model for curved DNA in which the A tracts are not curved at all, it seemed pretty heretical to many people in the field. Essentially the reasoning was that, since of the hundreds of DNA sequences that have been crystallized, A tracts are always rigid and never curved, the situation must be the same in solution.³³ This model proposes that the DNA curvature observed in gel electrophoresis experiments is due not to the curvature of A tracts, but to the comparative rigidity of the A tracts. The curvature arises not from tilt (remember that in the wedge model, the “wedges” consist of a combination of both tilt and roll), but almost exclusively from base pair roll, toward the major groove. The A tract has little roll (close to 0°); most base steps have, on average, a roll toward the major groove of about 3°, and certain sequences (notably GC) have a roll (again toward the major groove) of about 6° per step. Thus, it is the relative curvature of the intervening sequences flanking the A tracts that gives rise to the observed curvature, according to this model. This model predicts that the “opposite” curvature observed by Brukner for the GGGCCC curves, compared to the A tract curves, is relative to the average roll of about 3° found for most sequences. In this case the central GC roll of about 6° would be in the opposite direction (+3°) than for the A tracts (-3°), compared to most of the other steps.

D. A Tract DNA Adopts a Unique Double Helical Conformation

To better understand the DNA bending models, a bit more background about A tract DNA is needed. Part of the reasoning for the junction model was that A tract DNA could adopt a unique conformation, as discussed above. The general sequence model goes a bit further than the junction model, by saying that the A tracts are not responsible for curvature at all, but the idea that the unique B' helix of the A tract is somehow responsible in part for the observed curvature is the same. Over the years, much information has been obtained about A tract DNA.

1. *Rigidity of A Tracts*

Spectroscopic studies showed that short A tracts (of at least 4 bp) can adopt a structure similar to that of poly(dA)·poly(dT); this structure was independent of whether the A tracts were phased with the pitch of the helix.³⁴ The structure is somewhere between an A-DNA and B-DNA helix (sometimes referred to as a B'

helix), and has a rigid (and straight) structure, in which the base pairs have a high propeller twist (about 20°; B-DNA has a propeller twist near 0°). This conformation is stabilized by strong base-stacking interactions.³⁵ The AT base pair has a high propeller twist such that the A share an hydrogen bond with the T below, resulting in “bifurcated” hydrogen bonding.³⁶ The helical repeat of A tract DNA is only 10.0 bp/turn, in contrast to most other DNA sequences which have a helical repeat of about 10.5 bp/turn; this indicates that A tracts adopt a type of helix distinctly different from that of normal “generic” DNA.^{37,38}

2. *A Tracts Have a Narrow Minor Groove*

Burkhoff and Tullius³⁹ obtained results that were inconsistent with both the wedge and junction models. A hydroxyl radical footprinting technique was used to analyze the width of the minor groove. Since hydroxyl radical attacks ribose from the minor groove, the reactivity of DNA is dependent on the width of this groove. The hydroxyl radical footprinting pattern for B-DNA showed uniform cutting at each base. Kinetoplast bent DNA showed a reduction in cutting in the A tracts. This result was interpreted as reflecting a narrowing of the minor groove along the A tract.^{39,40} The polymers studied by Hagerman, $(GA_4T_4C)_n$ and $(GT_4A_4C)_n$, showed very different results. The hydroxyl cleavage pattern in the $(GA_4T_4C)_n$ bent polymer showed periodic reduction in the cleavage along the A tracts. There was no reduction in cleavage along A tracts in the $(GT_4A_4C)_n$ polymer. These results suggest that in the 5'-A₄T₄-3' sequence the run of A's can adopt a helix different from that of B-form DNA, a helix in which a gradual narrowing of the minor groove occurs. This narrowing does not occur in the A tracts in the 5'-T₄A₄-3' sequence. Probing with diethylpyrocarbonate also indicates that the minor groove is more narrow in A tract DNA; furthermore, addition of the minor-groove binder distamycin abolished a conformation of the A tracts that was KMnO₄ sensitive only at lower temperatures.⁴¹

The results from chemical probe analysis are consistent with the crystal structure of poly(dA)-poly(dT), in which the minor groove is more narrow than average B-DNA.³⁵ The narrowing of the minor groove for short A tracts associated with DNA curvature has also been found in crystal structures (see ref. 42 for a review) as well as in solution, based on nuclear magnetic resonance (NMR) experiments.^{43,44}

3. *Cooperativity in A Tract B' Helix Formation*

There is evidence for cooperativity in the formation of the A tract DNA structure. Leroy et al.⁴⁵ presented imino-proton-exchange NMR experiments to show that the structures of A_nT_n and T_nA_n sequences were quite different. Long proton exchange times for A·T base pairs were associated with longer lengths of A tracts in the A_nT_n but not the T_nA_n orientation. The shorter times in T_nA_n oligonucleotides were similar to lifetimes found in B-form DNA. Long proton exchange

times were found for all sequences that exhibited anomalous migration or polyacrylamide gels.⁴⁵ Nadeau and Crothers⁴³ have also confirmed that cooperative structural changes in helix structure occur in a run of A's as the length of the tract is increased. Three A's begin to set up the "A tract helix" responsible for the observed curvature. By the time a length of A₆₋₇ is reached, the transition from a B-form to a different helix, an A tract structure, is complete.

4. *Temperature Dependence of A Tract B' Helix*

Diekmann⁴⁶ showed a strong temperature dependence of curvature; he found that at 40°C, the curvature was greatly reduced. Curved sequences have been shown to undergo a "premelting" transition, with a midpoint near 30°C.⁴⁷ In fact, further studies have indicated that this premelting transition is due to the loss of a curved component from an equilibrium at lower temperatures.⁴⁸ This same premelting transition is seen for poly(dA)-poly(dT), which is known not to be curved, which means that the rigid high-propeller-twisted conformation of the A tracts is not directly responsible for curvature. The A tract regions from kinetoplast DNA were found to be sensitive to KMnO₄ at the 5' end of the A tracts, but only at lower temperatures—in this case there was a slight drop in reactivity at 23°C (compared to 14°C) and a dramatic reduction in reactivity at 43°C.⁴¹

It is certainly possible (indeed likely) that this conformation is indirectly responsible for the observed macroscopic curvature, by providing a frame of reference from which to see the curvature of the rest of the DNA. If this were true, one would expect the anomalous migration of A tract curved DNA in gels to be less significant above the transition temperature than below it, and this is exactly what is observed experimentally. There is little change in migration between 5°C and 25°C, but a drastic reduction in anomalous mobility if the gel is run at 35°C.

Obviously, such a temperature-dependent transition within this range could have strong biological implications. Many sequences that have been characterized at room temperature as being "curved" may in fact not be curved at all inside a cell at 37°C; furthermore, this difference in curvature could be utilized as a type of "environmental sensor." In support of this is the finding of large regions of significant curvature upstream of virulence-related genes in many different species of bacteria. It is possible that, when the bacteria are floating free in the environment, these genes are not transcribed, partly because of the curvature of the upstream region. However, at 37°C, when the bacteria have invaded a host, the upstream region is no longer as curved and might more readily allow transcription of the gene.

5. *Static versus Dynamic Curvature*

So far DNA curvature has been talked about as if it were a stable, static bend in the double helix. Molecular mechanics modeling of curved DNA shows a type of

flexing equilibria, such that on average a piece of DNA is curved, but it is in fact in equilibrium, or flexing, between two different conformations.⁴⁹ The models predict the flexing to be in the picosecond range. Thus the DNA appears, when averaged over time, to have a “static” curve of a given angle. This is consistent with NMR evidence, in which the “curved” conformation is due to flexing of the DNA flanking the A tracts (which are straight and have a high propeller twist).^{33,50}

6. Summary

Short A tracts can form a distinct type of structure, in which the helix is rigid, the minor groove is more narrow, and the bases are stacked with a high propeller twist. The formation of this structure is cooperative, with about six consecutive A’s needed for the most stable conformation to occur. This conformation is temperature dependent and appears to be stable at temperatures below 30°C. However, the presence of this structure does not always result in curvature, since the A₅ tract in the A₅N₁₀ sequence can also adopt this conformation, which is not curved. Furthermore, poly(dA)·poly(dT), which is also not curved, can adopt this conformation. The high propeller twist conformation, with the bifurcated hydrogen bonds, has been shown to be unnecessary for DNA curvature.⁵¹ Thus this structure is associated with curvature, but apparently is not the sole determinant.

E. Which Model Best Explains DNA Curvature?

The wedge and junction models have become refined, so that now they both give essentially the same predictions.⁵² Both of these models do allow for some curvature from non-A tracts (but the dominant contribution is assumed to come from phased A-tracts), and are supported by experiments using polyacrylamide gel electrophoresis (usually run in the presence of EDTA).

None of the models provide a completely accurate explanation of the physical basis for DNA curvature in solution. On the other hand, all of the models do a reasonable job of predicting the curvature (as determined by anomalous migration in polyacrylamide gels) of many DNA sequences containing runs of A’s. However, there are some sequences that exhibit curvature not predicted by either the wedge or junction models. Furthermore, the wedge and junction models do not readily account for the effects of Mg²⁺ and temperature on DNA curvature.

There is compelling evidence to support the general sequence model from many different physical methods. A comprehensive review of DNA curvature from 114 different crystal structures has found the A tracts to be “relatively straight,” with the curvature localized in flanking regions.⁴² In addition to the X-ray crystal structures, the gel electrophoresis experiments of G₃C₃ motifs,^{26,27} as well as sequence-specific flexibility, as determined by DNase I digestion patterns,^{53,54} lend credence to this model. This model is also consistent with studies in which the locations of particular dinucleotide steps were compared in DNA that was

wrapped around nucleosomes or into minicircles. The GC, CG, and GG steps were all found to prefer positions where the minor groove was on the outside of the circle, whereas the AA, AT, and TA dinucleotides were found to be positioned where the minor groove was facing inside^{55,56} (see Wolffe and Drew for a discussion).⁵⁷ Furthermore, theoretical modeling of curved DNA has also lent support to the general sequence model.⁵⁸ The thermodynamic evidence of a premelting transition in which the rigidity of the A tracts is lost also fits well with this model and is difficult to explain in terms of either the wedge or junction model.⁴⁸ Thus it seems that much evidence is currently pointing toward the model of DNA curvature in which the A tracts are not curved at all, most other steps are curved a bit because of a roll toward the major groove, and the GC step can exhibit strong curvature. However, one criticism of the conclusions from X-ray crystallography is the use of dehydrating agents and the frequent high concentrations of divalent cations needed for crystallography,⁵⁹ an issue that has been considered by Dickerson and colleagues.⁶⁰

The issue of the correct model for DNA bending remains to be resolved, and further experimental evidence will clearly be needed before the physical basis for bending is clearly resolved. It should be kept in mind that many assays for bending and DNA structure utilize very different protocols and procedures. For example, acrylamide gels can be run in the presence of many different buffers and in the absence of divalent cations, whereas circular ligation studies require Mg²⁺. X-ray crystallographic buffers are very different and contain dehydrating agents. Therefore, the assay conditions can influence the DNA structure, and this complicates direct comparison of results obtained using different methods. Finally, the environmental milieu surrounding the DNA in living cells is not well defined, adding another level of uncertainty to the structure of DNA in cells.

F. Environmental Influences on DNA Curvature

The environment surrounding DNA can have a profound influence on its structure. As has already been mentioned, at 37°C most A tract curved DNA will run on gels close to normal, and is likely to be hardly "curved" at all. Other factors known to affect DNA curvature are divalent cations, such as Mg²⁺, which will enhance the magnitude of curvature of many sequences. On the other hand, addition of NaCl to about 300 mM will greatly reduce DNA curvature. The addition of spermine (or spermidine), which is abundant in most cells, will again enhance curvature. Supercoiling of the DNA will also affect the manifestations of local curvature. Of course, the binding of a protein to DNA will have a large effect on the curvature, but this is DNA bending, the subject of the next section.

G. Proteins That Bind and Bend DNA

The DNA inside of most living organisms is compacted at least 1000-fold, and it may not be surprising that just about every DNA-binding protein will bend or

wrap the DNA around itself when it binds. There are many proteins that will bind preferentially to DNA with intrinsic curvature, as well as proteins that bind to DNA and then bend it. This is something of an artificial distinction, in that most of the proteins that bind preferentially to curved DNA will, in all likelihood, bend the DNA even more. Before the proteins themselves are discussed, it is important to distinguish between DNA sequences that are intrinsically curved and sequences that are flexible.

1. Intrinsic Curvature versus Flexibility

“Flexible” sequences contain steps (usually pyrimidine-purine, such as TA or CA) that are easily deformable.⁶¹ When a curved DNA fragment runs anomalously slow on the gel, there are two components to the observed “curvature”: one is the actual “static” or intrinsic curvature, as has been discussed so far in this section. The other component is the “bendability” or flexibility of the DNA sequence. Some DNA sequences are quite readily deformable, whereas others (like the A tracts) are quite rigid. In fact, one aspect of the general sequence model for DNA curvature is that, because the A tracts are rigid and do not bend, the comparative flexibility of the adjacent sequences contributes to the observed curvature. In some crystal structures, a motif has been found to be bent in one sequence context, but not in another; this has been explained as “a demonstration of the bendability of the helix.”⁶⁰

The bendability of various DNA sequences can also be determined by locating the relative positions of dinucleotides in DNA wrapped around nucleosomes or in tight minicircles.^{55,56} The result of several such studies is that GC dinucleotides are found preferentially where the minor groove is facing out, or away from the protein–DNA complex, whereas the AA dinucleotide was found preferentially in the opposite orientation—that is, with the minor groove facing in, or toward the center of the complex.⁵⁷ Thus the flexibility of the GC step will allow it to roll more toward the major groove.

2. Proteins That Bind to Curved DNA

Proteins that bind nonspecifically to DNA, yet show a preference for sequences containing curved DNA, are often chromatin-associated proteins and are responsible for maintaining the chromosome. Travers has called these proteins “DNA chaperones,” since one of their main roles seems to be facilitating DNA compaction and other protein–DNA interactions.^{62,63} Three different categories will be considered: the bacterial proteins, eukaryotic “curved DNA-binding” proteins, and eukaryotic nucleosomes.

Curved DNA-binding proteins in bacteria. Both of the major chromatin-associated (HU and H-NS) proteins in *Escherichia coli*, as well as the *E. coli* RNA

polymerase, have been found to bind preferentially to curved DNA, with the H-NS protein showing the greatest affinity for curves.⁶⁴ The H-NS protein was isolated from *E. coli* cell extracts, based on its preference for curved DNA.⁶⁵ H-NS binding to DNA is sensitive to intercalating agents, such as distamycin,^{65,66} which will remove intrinsic curves from DNA.⁶⁷

Curved DNA sequences seem to occur within the sites of H-NS interaction in genes repressed by H-NS—at least in the cases of three H-NS dependent genes studied to date: the *proU* operon of *Salmonella*,⁶⁸ the *rrnB* gene of *E. coli*,⁶⁶ and the *hns* gene itself.^{69–71} Furthermore, *in vivo* experiments seem to imply that H-NS can show strong and specific repression of a gene with a curve upstream of the promoter region.⁷² It is worth noting in this last experiment that a different curve resulted in no repression by H-NS *in vivo*. The surprising thing about this was that this curve (A₅N₅) was the same curve that had been used to identify H-NS as a “curve binding” protein.^{65,73} Thus it is possible that H-NS might show a strong binding preference for a particular type of curve, but exhibit different properties in binding DNA inside the cell. In fact, there are many curves (both synthetic and naturally occurring) that do not show a preference for H-NS binding. Upstream of the *E. coli proU* operon, there is a strong curve region, as found by random cloning of curved DNA fragments.⁷⁴ Although H-NS affects the regulation of the *proU* gene, it appears this occurs through a region downstream of the promoter, and deletion of the upstream curve has essentially no effect on H-NS regulation.⁷⁵ In summary, there is a correlation between the sites of H-NS interaction and DNA curvature, but not all curves will necessarily show a strong preference for binding of the H-NS protein.

Curved DNA-binding proteins in eukaryotes. Several eukaryotic proteins have been reported to bind preferentially to curved DNA sequences. One such protein is the mouse Kin17 protein.⁷⁶ The domain of the protein responsible for binding to curves is not known, although mutational analysis has shown that it does not act through the zinc finger region.⁷⁶ A protein that binds specifically to rat mitochondrial DNA has been shown to exhibit part of this preference through binding to curved DNA.⁷⁷ Finally, the chromatin associated HMG proteins exhibit an “architecture preference”—that is, they will bind to DNA of a particular type of flexibility.⁷⁸

Nucleosomes. Nucleosomes can be “phased” or positioned by certain sequences containing curved DNA.^{74,79} However, attempts to phase nucleosomes using synthetic curves so far have been unsuccessful.⁶⁴ It is likely that the phasing of nucleosomes will require many elements, including properly positioned flexible regions and flanking regions of “straight” DNA (see Wolffe and Drew⁵⁷ for a more detailed discussion). Other alternative DNA structures can also affect nucleosome position (see Section IV).

3. Protein-Induced DNA Bending

There are many examples of DNA bending upon protein association. In many cases, the DNA is actually wrapped around a protein core—for example, the wrapping of DNA around DNA gyrase, nucleosomes, or the bacterial HU protein. Many other proteins bind to DNA and introduce a bend that is similar in magnitude to the curvature introduced by A tracts in DNA. The catabolite activator protein (CAP) provides one example. Wu and Crothers,²³ using gel mobility and circular permutation analysis (see above), demonstrated that CAP binding introduces a significant bend into DNA. The X-ray cocrystal structure of the CAP-DNA complex has subsequently shown a bend angle of 90°.⁸⁰

H. The Biology of DNA Curvature

Nature has found many uses for DNA curvature. One is the control of access to promoters, the switch regions that turn genes on by bending or looping the DNA. Another is the control of initiation of DNA replication, which, since it represents a major commitment for the cell, must be very carefully regulated. A third use is site-specific integration of one DNA molecule into another. A fourth use is in DNA repair, where DNA often becomes bent upon the binding of many chemicals or following UV irradiation. Finally, the compacted organization of DNA into cells requires DNA to be wrapped very tightly around DNA-packaging proteins.

1. DNA Curvature, Bending, and Gene Regulation

DNA bending and open-complex formation. The RNA polymerase must bind and bend the DNA, and this bending (and subsequent torquing) is responsible for melting the DNA and the formation of an open complex.^{81,82} The minor groove at the center of the -10 site must be placed on the inside of an intrinsic curve for optimal promoter activity.^{55,83,84} Mutations that affect the curvature at the -10 site also affect the sensitivity of the promoter to DNA supercoiling.⁸⁵

Location of DNA curves in promoter regions. Much of the curved DNA in *E. coli* has been found to be localized in and upstream of promoter regions.^{74,86} The curved regions around promoters can be grouped into three different areas: upstream, downstream, or within the promoter.

Upstream curves. Upstream curves can have a strong influence on transcription, as has been demonstrated in many studies.^{87–89} A comparison of 43 different promoters in *E. coli* found a strong correlation between the presence of upstream curves and promoter strength.⁹⁰ The promoter region of the β-lactamase gene of plasmid pUC19 has an upstream curved region, and although no effect of changing the spacing of the curve was seen on transcription *in vitro*, there was a strong

dependence on rotational phasing of the curve with respect to the promoter *in vivo*.⁹¹ This particular curve was stable, even at 60°C, in contrast to most A + T-rich curves, which often exhibit little curvature, even at 40°C, as discussed above.⁴⁶ Kim et al. have recently shown that upstream curves can enhance the effects of transcription factors.⁹² There are many examples of upstream curves influencing transcription,⁸⁶ but perhaps one of the most striking examples of the importance that DNA structure can play was the experiments by Goodman and Nash, in which they replaced the binding site of a protein that was known to bend DNA, with a synthetic curve of approximately the same magnitude, and showed function in the absence of the protein.⁹³

Tanaka et al. have found that the curve upstream of the *proU* promoter must be positioned properly to get high levels of reporter gene activity.⁹⁴ There was a very strong correlation between the relative gel mobility of the spacer inserts (reflecting the orientation between the upstream curve and the promoter) and β-galactosidase expression, which varied about 100-fold, in a "face of the helix" manner, depending on the spacing of the insert.⁹⁴

Curves within the promoter region. Within the promoter region itself, curved DNA plays an important, but perhaps bit more subtle role. One of the most obvious functions is the formation of a proper RNA polymerase binding site. Ross et al. have shown that there is a "third recognition element" in some bacterial promoters, called the UP element, which is located between -60 and -40; this interacts with the C-terminal of the α-subunit of RNA polymerase.^{95,96} DNA "curvature" is important, at least in the sense of relative orientation of this region of DNA with the α-subunit of RNA polymerase. The consensus sequence is A + T rich, which could reflect the importance of flexibility in bending or wrapping the DNA about the protein.⁹⁵

In addition to this UP element, curves within the spacer element between the -35 and -10 sites can have a determinant role in promoter strength.^{84,89} The geometry and flexibility of the promoter are important in determining efficient binding and initiation of RNA polymerase. For example, it has been suggested recently that the suboptimal spacing between the -35 and -10 regions of the *proU* promoter must be compensated for by an increase in flexibility with this spacer region.⁹⁷ This increased flexibility allows the -35 and -10 regions of this promoter to adopt a more favorable position for binding of the RNA polymerase σ-factor.

Downstream curves. Most endogenous curves have been found upstream or within the promoter regions.⁷⁴ In comparison, only a few curves have been characterized downstream of the transcription start site. The few reported examples of curves downstream from promoters seem to have properties distinctly different from those of the upstream curves, perhaps reflecting different roles in transcription. For example, Schroth et al. have found that there are two distinct curves flanking the promoter of an rRNA gene from *Physarum polycephalum*; one curve

is centered roughly 160 bp upstream, and the other is about 150 bp downstream of the transcription start site. Although both are curved by the same magnitude (approximately 45°), the curves exhibited different behaviors in terms of temperature dependence of the curve, the effects of EtBr, and relative mobility when the percentage of acrylamide was varied.⁷⁹ Thus these two curves would be expected to change in different ways under various environmental conditions.

III. STRUCTURE AND FUNCTION OF SUPERCOILED DNA

A. Introduction

DNA normally exists in a supercoiled form in most biological systems. Supercoiling makes structural variations in shape and helix structure particularly dynamic, with a wide variety of conformations possible through variations in twisting and writhing. Supercoiling appears to be very critical for viability in bacterial systems. *E. coli*, for example, can only tolerate mutations that change supercoiling levels by less than 25%, otherwise lethality ensues. Supercoiling is important for gene expression, DNA replication, and recombination. Although supercoiling exists in mammalian genomes, its biological significance is not yet appreciated.

Historically, various “forms” of DNA molecules have been identified, representing different topological conformations. A plasmid molecule purified from a bacterial cell will exist as a naturally occurring covalently closed circular DNA molecule that is negatively supercoiled, historically called form I DNA. DNA containing a single nick in one of the strands will lose all supercoils. The nicked molecule (also called open circular DNA) is called form II DNA. DNA that contains breaks in both phosphate backbones at the same point (or nearly the same point) along the helical axis will form a linear DNA molecule, known as form III. Form IV refers to denatured DNA (and closed circular catenated). Form V DNA is formed when two (not catenated) circular single strands anneal. Half of the DNA can form right-handed B-form DNA, and for topological reasons, half of the DNA must form left-handed turns.

B. Supercoiled Forms of DNA

1. $Lk = Tw + Wr$

Supercoiled DNA is characterized by a topological property called the linking number, Lk . Lk is defined as the number of times one strand crosses the other when oriented in a plane. Lk must be an integer. The relaxed DNA molecule shown in Figure 12A has a linking number $Lk = 20$. The linking number can only change when the phosphodiester backbone is broken by chemical or enzymatic cleavage.

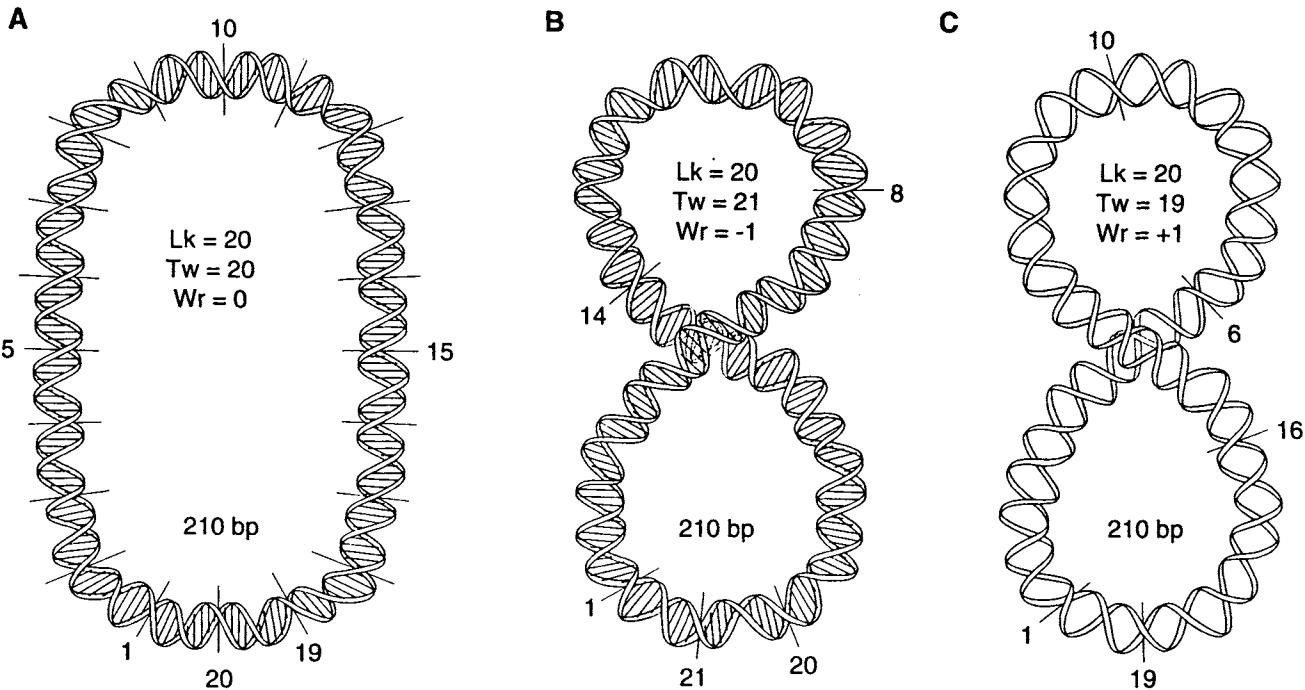


Figure 12. Positive and negative supercoils in DNA. (A) A covalently closed, relaxed DNA molecule containing 210 bp of DNA is shown. This molecule contains 20 helical turns, with $L_k = 20$, $T_w = 20$, $W_r = 0$. (B) This molecule contains one right-handed supertwist or writhe ($W_r = -1$; a negative supercoil). Because L_k cannot change, the twist must change by +1 ($T_w = 21$). (C) This molecule contains one left-handed supertwist (or positive supercoil; $W_r = +1$), and consequently $T_w = 19$.

In a covalently closed molecule, although Lk cannot change, the number of twists or turns of the double helix, as well as the number of supercoils or writhes of the double helix, can change. DNA topology is described by the simple equation

$$Lk = Tw + Wr \quad (1)$$

where Lk is the linking number, Tw is the number of helical turns in the DNA, and Wr is the writhing number of DNA. Wr describes the supertwisting or coiling of the helix in space.

The introduction into a DNA molecule with $Lk = 20$ of a negative supertwist that is a right-handed coil changes the value of Wr by -1 ($Wr = -1$) (Figure 12B). Since $Lk = 20$ and Lk cannot change, Tw must increase by +1 to a value of 21. Conversely, a decrease in the twist number by -1 to a value of 19 would require a compensating introduction of a left-handed or positive supertwist ($Wr = +1$), as shown in Figure 12C.

2. Relaxed DNA

On average, the helical repeat of DNA is about 10.5 bp per helical turn of DNA. In linear or nicked DNA, where the ends of the molecule are free to rotate, the DNA will adopt a preferred helical repeat. The preferred helical repeat of a nicked or linear DNA molecule represents the lowest energy form of the molecule. When this state of helical twist exists in a covalently closed molecule, the molecule is relaxed and contains no supercoils. In relaxed DNA, as shown in Figure 12A, the linking number equals the twist number ($Lk = Tw = 20$ and $Wr = 0$). The linking number of relaxed DNA, Lk_0 , is defined as

$$Lk_0 = N/10.5 \quad (2)$$

where N is the number of base pairs in the DNA molecule and 10.5 refers to the helical repeat.

3. Negatively Supercoiled DNA

Negatively supercoiled DNA has a deficiency in the linking number compared to relaxed DNA, or $Lk < Lk_0$. Negatively supercoiled DNA is underwound with respect to helical turns, i.e., it contains fewer helical turns than the molecule would contain as a linear or relaxed molecule. This underwinding in the number of helical turns results in more base pairs per helical turn compared to B-DNA and in a decrease in the angle of twist (or the rotation/residue) between adjacent base pairs. This underwinding creates torsional tension in the winding of the DNA double helix.

The topology of negatively supercoiled DNA is illustrated in Figure 13. In Figure 13A, the 210-bp relaxed molecule has 20 helical turns, and $Lk_0 = Lk = 20$. In

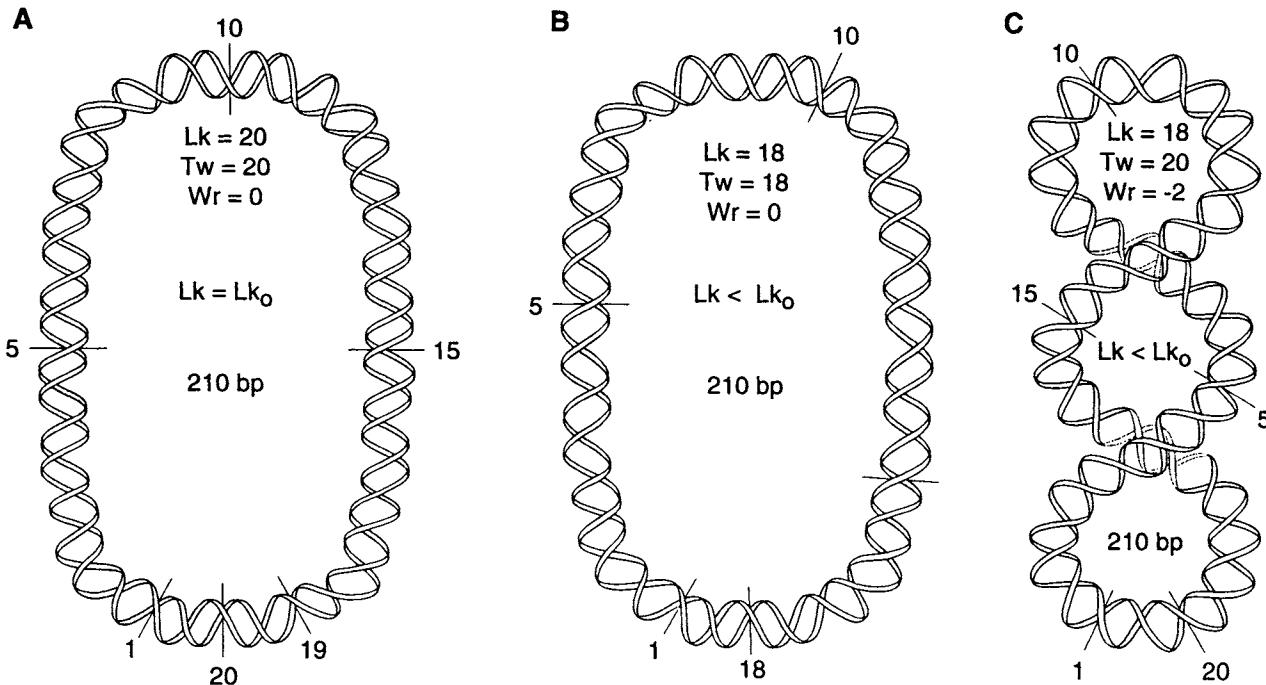


Figure 13. Introduction of negative and positive supercoils into DNA. (A) In this relaxed DNA molecule $Lk = 20$, $Tw = 20$, $Wr = 0$. (B) In this negatively supercoiled DNA molecule, the linking number deficit is 2 ($Lk = 18$), resulting in an increase in the number of base pairs per turn, a decreased angle of rotation, and $Lk < Lk_0$. (C) In this molecule two negative supercoils are introduced into the DNA such that $Lk = 18$, $Tw = 20$, $Wr = -2$.

Figure 13B, with $Lk = 18$, this 210-bp DNA has only 18 helical turns. This will change the average rotation per residue from 34.29° [$(20 \times 360^\circ) / 210$] to 30.86° [$18 \times 360^\circ / 210$], which represents an unfavorable winding of the DNA double helix.

The molecule in Figure 13B is negatively supercoiled, since $Lk < Lk_0$, but it does not contain the familiar interwound supercoil in the molecule of Figure 13C. The linking number deficit is manifested as a twist deficit in the molecule shown in Figure 13B. It is the torsional strain in the winding of the DNA helix, a result of the decreased twist angle of DNA, which drives the supertwisting of the DNA into the supercoils shown in Figure 13C. Negatively supercoiled DNA forms a right-handed (or clockwise) supercoil. Since Lk_0 remains the same ($Lk_0 = 20$), the DNA is supercoiled by the amount $\Delta Lk = (Lk - Lk_0) = -2$. In winding in these two negative supercoils ($\Delta Wr = -2$), two additional helical turns are wound into the helix ($\Delta Tw = +2$).

4. Positively Supercoiled DNA

Although most DNA isolated from natural sources is negatively supercoiled, DNA can exist in a positively supercoiled form. DNA is said to be positively supercoiled when $Lk > Lk_0$. Positively supercoiled DNA is overwound in terms of the number of helical turns, resulting in fewer base pairs per helical turn and an increase in the winding angle between adjacent base pairs.

Overwinding a 210-bp DNA molecule by two turns to $Lk = 22$ creates the average rotation per residue from 34.29° (in relaxed DNA) to 37.71° [$(22 \times 360^\circ) / 210$] (Figure 13). This, like the situation for negatively supercoiled DNA, represents an unfavorable state. The tension in the winding of the helix is relieved by the positive supercoiling of the DNA forming a left-handed (counter clockwise) supercoil (Figure 13D). As two positive supercoils form ($\Delta Wr = +2$), two helical turns are removed returning the number of helical turns to $Tw = 20$, which is the preferred conformation for the DNA double helix.

Positively supercoiled DNA has been isolated in what appears to be a naturally occurring form. A bacteriophage-like plasmid molecule from a *Sulfolobus* species, an archebacterium living at high temperature and low pH, has been shown to contain positive supercoils.⁹⁸ Positive supercoiled DNA would resist unwinding of the helix by heat and acid. Packaging DNA in a positively supercoiled form may be one mechanism for protecting the genetic information from denaturation.

5. Supercoils: Interwound or Toroidal Coils

In addition to existing as interwound supercoils, negative supercoils can exist as left-handed toroidal coils (Figure 14). Toroidal coils topologically satisfy the requirement for Wr , although in a toroidal coil the helix does not cross itself in a plane, as is the situation for interwound supercoil. The DNA duplex does cross

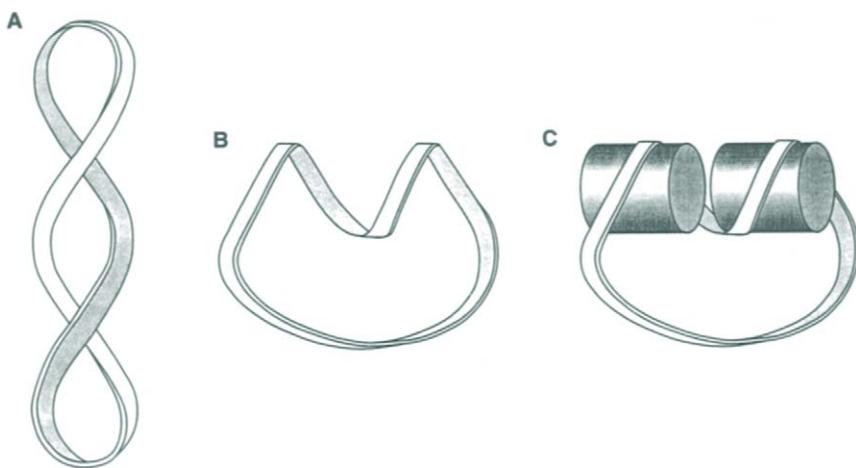


Figure 14. Right-handed interwound negative supercoils and left-handed toroidal coils. (A) A molecule with two interwound negative right-handed supercoils. (B) A molecule with two left-handed toroidal coils. (C) Two toroidal coils can be wrapped around a protein (represented by the cylinders).

itself in the plane of the toroidal coil. In solution, supercoils are distributed in part by a decreased angle of twist and a mixture of interwound and toroidal coils.

Knots and catenanes can have a negative or positive sign, depending on the order of the strand crossings (also called nodes). The convention for determining the sign of nodes is described in the legend to Figure 15. DNA knots and catenanes are found *in vivo* as the product of certain topoisomerases and site-specific recombination enzymes.^{99–101}

6. Superhelical Density and the Specific Linking Difference of DNA

Two terms, *superhelical density* (σ) and *specific linking difference* (σ_{sp}), are frequently used to describe a level of supercoiling. Superhelical density, σ , is defined as the average number of superhelical turns per helical turn of DNA:

$$\sigma = 10.5 \tau/N \quad (3)$$

where τ is defined as the titratable number of superhelical turns, N is the number of base pairs in the molecule, and 10.5 represents the average number of base pairs per turn. τ rather than Wr is used for historical reasons (see ref. 1). Since there are now many ways to measure the number of supercoils (electron microscopy, two dimensional agarose gels, as well as the original solution titration methods), τ will be defined as the number of measurable supercoils.

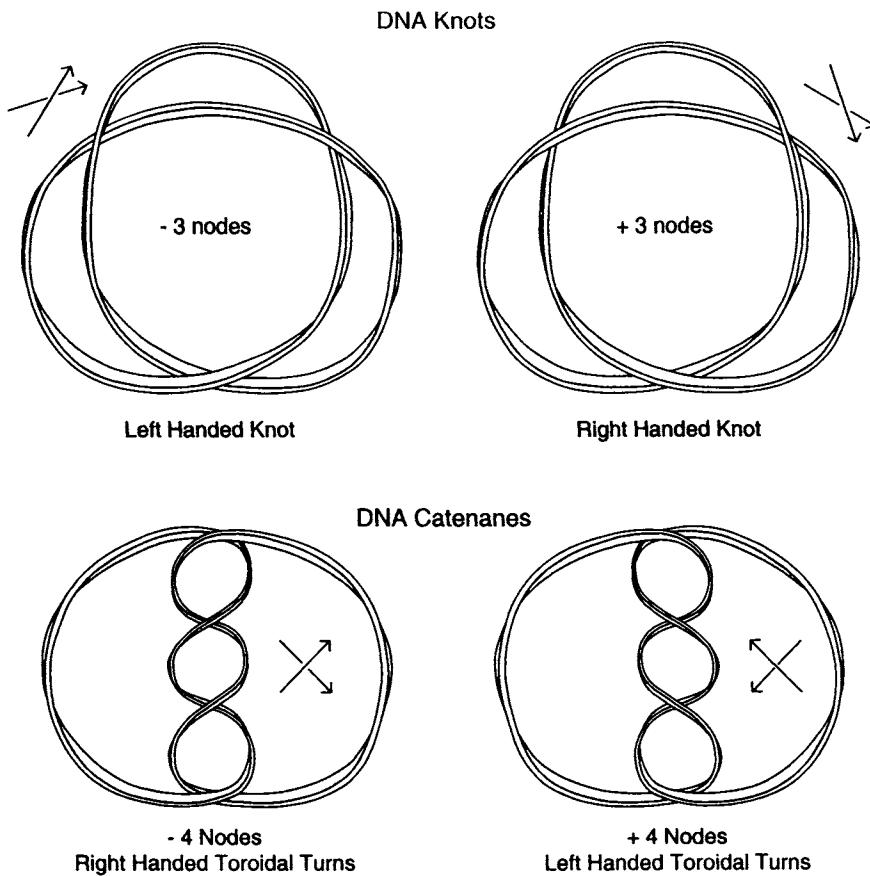


Figure 15. DNA knots and catenanes. (**Top**) Shown are a left-handed and a right-handed knot with negative and positive nodes, respectively. By convention, a negative node occurs when an arrow drawn along the top strand is rotated $<180^\circ$ in a clockwise direction to align with an arrow drawn in the same direction along the bottom double helix. When the top arrow must be rotated counterclockwise $<180^\circ$ to align with the bottom arrow, the node is positive. (**Bottom**) Shown are left-handed and right-handed antiparallel catenanes, each with four nodes. The positive nodes are formed by left-handed (or negatively supercoiled) toroidal turns, whereas the negative nodes are formed by right-handed (or positively supercoiled) toroidal turns.

The specific linking number difference, σ_{sp} , a term that is also used to describe the level of superhelical density, is defined as

$$\sigma_{sp} = (Lk - Lk_0)/Lk_0 \quad (4)$$

7. *The Energetics of Supercoiled DNA*

The free energy of supercoiling is proportional to the square of the linking number difference according to the following relationship:

$$\Delta G = (1100 RT/N)(Lk - Lk_0)^2 \quad (5)$$

where R is the gas constant, T is the temperature in degrees Kelvin, and N is the number of base pairs in the DNA molecule. The free energy in supercoiled DNA can be used to drive biological reactions. Transcription and DNA replication require energy to open or unwind the DNA double helix to allow access of the bases at the center of the DNA helix. Some of this energy comes from supercoiling.

8. *Effects of Temperature on DNA Supercoiling*

The twist of the DNA double helix is sensitive to changes in temperature. As temperature decreases, the twist angle in DNA increases, resulting in fewer base pairs per helical turn. This results in an increase in Lk_0 . In a covalently closed DNA molecule where Lk cannot change, the temperature induced change in helical winding changes Lk_0 , which concomitantly affects $Lk - Lk_0$, or the superhelical density. The temperature induced change in Lk_0 is $\Delta Lk = 0.012$ helical degrees ($^{\circ}$)/bp/ $^{\circ}$ C.^{102,103}

C. The Biology of Supercoiled DNA

1. *Topological Domains of DNA: A Requirement for Supercoiling*

Most DNA, including bacterial chromosomes and plasmid, as well as linear chromosomal DNA and small circular molecules in human cells, is negatively supercoiled. Covalently continuous circular molecules are topologically closed and thus have a defined linking number. DNA molecules must exist in a topologically closed form to be supercoiled. The *E. coli* chromosome exists as a large (2.9×10^6 bp) closed circle that is subdivided into 45 ± 10 independent topological domains *in vivo*. A topological domain is defined as region of DNA bounded by constraints on the rotation of the DNA double helix. The nature of the molecules responsible for the topological domains *in vivo* remains to be established. Domains may be created by the attachment of DNA to the bacterial membrane through specialized binding proteins, DNA gyrase, or other topoisomerases (such as Topo IV).

Linear DNA must become organized into one or several topological domains to exist in a supercoiled state. Following infection, the ends of the linear 120-kb bacteriophage T4 chromosome are prevented from rotating, thus organizing the entire chromosome into a single topological domain. The model for the organization of

human (eukaryotic) chromosomes involves independent loops, likely formed by the interaction of specific regions of DNA with defined proteins that attach to the nuclear matrix.

2. DNA Topoisomerases

Topoisomerases transiently break and reseal phosphodiester bonds in DNA, allowing one strand to pass through the other strand. This effectively provides a swivel to allow one strand to rotate around the other. It is topoisomerases that are responsible for changes in the linking number of DNA in living cells. In bacterial cells, negative DNA supercoils are introduced by DNA gyrase, and they are removed by DNA topoisomerase I. These enzymes act by breaking the phosphodiester bond and forming a covalent bond between a specific tyrosine in the protein and a phosphate group on the 5' or 3' end of the DNA. In some cases the enzymes require ATP as an energy source (see Sinden¹ for details and specific properties).

Type I and type II topoisomerases. Type I topoisomerases break one strand of the DNA, and type II topoisomerases break both strands of the DNA. The linking number of DNA, Lk , will correspondingly change in increments of 1 or 2 for type I and type II topoisomerases, respectively. *E. coli* topoisomerase I, the first DNA topoisomerase to be discovered, relaxes negatively supercoiled DNA. Eukaryotic type I topoisomerases can relax negative or positive supercoils. ATP is not required for type I topoisomerase activity. DNA gyrases, type II topoisomerases, are the only enzymes that can catalytically introduce negative supercoils into relaxed DNA in an ATP-dependent process. In the absence of ATP, gyrase can relax DNA.

The biological importance of DNA topoisomerases. Topoisomerases maintain a precise level of supercoiling in *E. coli*. Mutations in DNA gyrase can result in a decrease in the level of supercoiling in living cells. Similarly, mutations in topoisomerase I can result in an increase in the level of supercoiling in cells. The topoisomerase genes encoding DNA gyrase and topoisomerase I are essential. The cell will only tolerate mutation resulting in changes in linking number of about 25%. A deletion mutation of topoisomerase I was viable when a compensatory mutation in DNA gyrase occurred, reducing the ability of this enzyme to supercoil DNA. A number of excellent reviews on DNA supercoiling and the regulation and genetics of the level of supercoiling in *E. coli* are available.^{104–109}

All biological reactions involving DNA likely require DNA topoisomerases. For example, following replication, two circular chromosomes will become catenated. A type II topoisomerase activity is required to transiently break the DNA double helix, allowing a duplex from the second chromosome to pass through the duplex of the first chromosome. Without this strand-passing reaction following replication, the two chromosomes cannot physically separate and the cell cannot divide.

As discussed later in this section, the movement of RNA polymerase can generate superhelical tension. Topoisomerases are needed to relax positive and negative supercoiling generated by proteins tracking through the DNA. Without these enzymes, transcription and replication would likely cease.

3. *Mechanisms of Supercoiling in Cells*

Supercoiling in bacterial cells is driven by the activity of DNA gyrase. DNA gyrase was the first enzyme identified that can negatively supercoil DNA. Purified DNA gyrase introduces negative supercoils to a superhelical density of about $\sigma = -0.1$ in a reaction requiring ATP hydrolysis. In living cells superhelical density is regulated by the opposing activities of DNA gyrase and topoisomerase I. The balance of supercoiling by DNA gyrase and relaxation by Topo I keeps the DNA at a finely tuned level of supercoiling in living cells. The ATP/ADP ratio is also very important in maintaining the level of supercoiling *in vivo* as gyrase activity is influenced by this ratio.¹¹⁰

Supercoiling in eukaryotic cells may be introduced by organization of DNA into nucleosomes and topoisomerase activity. In eukaryotic cells, DNA is wrapped in a left-handed fashion around a histone octamer, forming a nucleosome. The left-handed toroidal coiling of DNA around a protein, which restrains a negative supercoil, introduces a positive supercoil into the nonnucleosomal DNA in the same topological domain, with no change in Lk . Either eukaryotic type I or type II topoisomerases can relax the positive supercoil, resulting in the net introduction of a negative supercoil. Note that wrapping DNA twice around the nucleosome ($\Delta Wr = -2$) results in the introduction of a single positive supercoil. This linking number paradox, where $\Delta Lk = -1$ and $\Delta Wr = -2$, can be resolved by a compensatory change in Tw (equal $\Delta Tw = +1$). The helix repeat changes from about 10.5 to <10.5 in DNA that is wrapped around a nucleosome.

Transcriptional effects on supercoiling. In 1986, Pruss and Drlica discovered that the negative superhelical density of a plasmid was twice that normally found *in vivo* when the plasmid was purified from cells containing a mutation in topoisomerase I.¹¹¹ The plasmid sequences responsible for this unusually high supercoiling included the promoter region for the tetracycline resistance gene. Moreover, the unusually high supercoiling was dependent on transcription from the tetracycline gene. In 1983, Lockshon and Morris showed that positively supercoiled plasmid DNA could be purified from *in vivo* cells treated with novobiocin, an antibiotic that inhibits the activity of DNA gyrase.¹¹² Liu and Wang then argued that the movement of an RNA polymerase during transcription would generate negatively supercoiled DNA behind the RNA polymerase while generating positively supercoiled DNA in front of the RNA polymerase (Figure 16).¹¹³ Within a topological domain of DNA when RNA polymerase moves through the

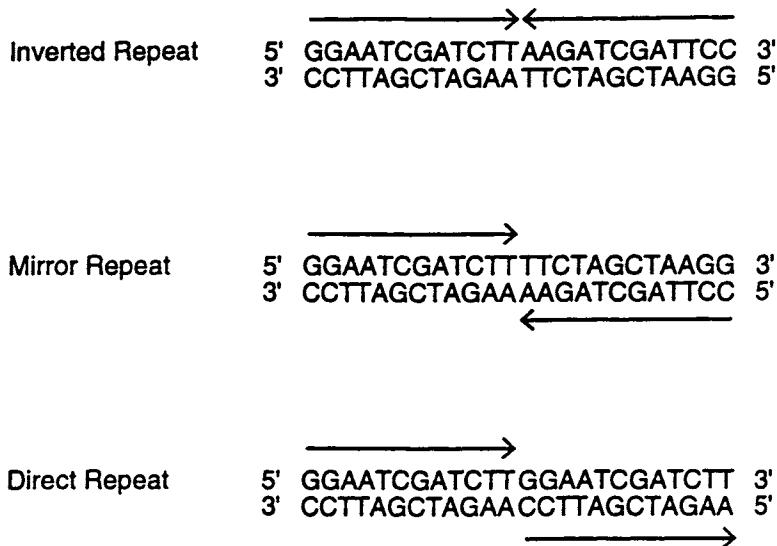


Figure 16. Transcription-dependent DNA supercoiling. Transcription can introduce DNA supercoiling as described in the text. (A) The single topological domain, which contains 210 bp of DNA, is divided into two subdomains by an RNA polymerase molecule. (B) Movement of RNA polymerase by 20 bp without rotation of the RNA polymerase results in a decrease in helical turns ($Lk < Lk_0$) behind the polymerase and an increase in helical turns ($Lk > Lk_0$) in front of the polymerase.

DNA without rotation, it divides the domain into two topologically separate sub-domains. The movement of RNA polymerase (when Lk is invariant) changes the relationship of $Lk - Lk_0$, since Lk_0 behind RNA polymerase increases, whereas Lk_0 in front of RNA polymerase decreases. This creates two subdomains in which $Lk < Lk_0$ behind the polymerase and $Lk > Lk_0$ in front of the polymerase, introducing negative and positive supercoiling, respectively. To introduce negative supercoiling into DNA during transcription, topoisomerase activity must relax the positive supercoils ahead of RNA polymerase.

DNA supercoiling and gene expression. The state of DNA supercoiling can influence the regulation of gene expression, as reviewed by Drlica,¹⁰⁴ Esposito and Sinden,¹¹⁴ and Freeman and Garrard.¹¹⁵ The conformation of DNA can also influence gene expression. The very different structure of cruciforms, Z-DNA, or triple-stranded regions might provide a switch that could turn a gene on or off. A protein, for example, might bind to a B-form DNA sequence to turn on a gene, and

it could not bind if the sequence existed in the alternative configuration. A model system using an inverted repeat that can form a cruciform has been shown to modulate gene expression *in vivo*.¹¹⁶

The energy from DNA supercoiling can be used to drive the opening of the promoter region, thus facilitating RNA polymerase binding. Different promoters are “tuned” for optimal transcription at different levels of supercoiling. This can be accomplished in a number of ways. For example, changes in the A + T content of the promoter could influence melting. The spacing of the -35 and -10 regions allows enormous variation in the promoter characteristics. Supercoiling differentially affects five or six kinetically distinct steps in transcription initiation, as reviewed by McClure.¹¹⁷

The regulation of the *E. coli* DNA gyrase and topoisomerase genes are a classic example of DNA supercoiling-dependent gene regulation. Transcription from the DNA gyrase gene is turned on by a low level of supercoiling and is turned off when the level of supercoiling is high. Moreover, the topoisomerase I gene is turned on when supercoiling levels are high. Thus while the opposing enzymatic activities of DNA gyrase and topoisomerase I regulate the level of supercoiling, the levels of these enzymes are also regulated by DNA supercoiling.

A supercoiled DNA template is required for precise initiation of DNA replication at the *E. coli* DNA *oriC*. The *oriC* DNA is wrapped into a toroidal coil around a complex of DnaA proteins. In an ATP-dependent process, three 13-bp AT-rich direct repeats are stably unwound, exposing this region to recognition and binding by a DnaB–DnaC complex. Other proteins then assemble on DNA to initiate replication from *oriC*.

IV. CRUCIFORM STRUCTURES

A. Introduction

Within the DNA there are defined, ordered sequences (dsDNA). Such elements include inverted repeats, mirror repeats, and direct repeats (Figure 17), which can form cruciforms, intramolecular triplex DNA, and slipped mispaired structures, respectively.

An inverted repeat (or palindrome) is a DNA sequence that “reads” the same in either direction (from the 5' to 3' in either strand). Inverted repeat sequences are widely distributed in the chromosomal DNA of many eukaryotes, including plants, yeast, *Neurospora*, *Physarum*, *Drosophila*, mouse, *Xenopus*, and human (for a review, Pearson et al.).¹¹⁸ In all cases their distribution is nonrandom and is clustered at or near regions of genetic regulation. Cruciform structures may transiently form under physiological conditions to serve as recognition signals for regulatory proteins of transcription, recombination, or replication (see above). Some possible roles that inverted repeats may play are described below.

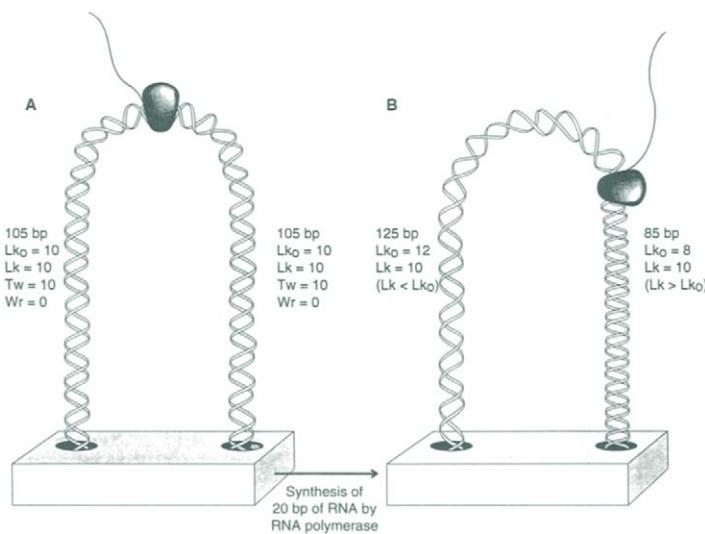


Figure 17. Defined ordered sequence DNA (dosDNA). Inverted repeats, mirror repeats, and direct repeats represent defined ordered sequence DNAs (dosDNAs). The arrows above and below the base sequences show the organization of symmetrical complementary sequences in DNA. In inverted repeats (or palindromes) the two strands of the DNA are complementary to each other, as well as being self complementary. A mirror repeat has identical base pairs in one strand surrounding a center of symmetry. Arrows show complementary base pairs in the mirror repeat sequence. Certain mirror repeats can form intramolecular triplex structures (as described in Section VI). Direct repeats contain a particular sequence that is repeated or duplicated. Direct repeats can be adjacent (as in the example shown), or they can flank an intervening sequence. Direct repeats can form slip-mispaired structures (as described in Section VII).

B. Formation and Stability of Cruciforms

1. Cruciform Formation Requires DNA Supercoiling

Energy from DNA supercoiling is involved in melting the center of the inverted repeat, allowing the intrastrand nucleation required for cruciform formation. For cruciform formation about 10 bp must unwind at the center of symmetry (Figure 18A).^{119–121} This provides a region in which intrastrand base pairing can occur. Following nucleation, the inverted repeat extrudes as a cruciform. Cruciforms form in supercoiled but not relaxed DNA.^{122–124} A very specific minimal level of superhelical energy is required for cruciform formation, σ_c .¹²⁵ DNA supercoiling is also required for cruciform stability. Stability comes from the relaxation of negative supercoils as cruciforms are formed. Cruciform formation results in the relaxation of one negative supercoil per 10.5 bp of DNA that forms the cruciform. The relaxation of supercoils reduces the free energy (ΔG) of supercoiling.

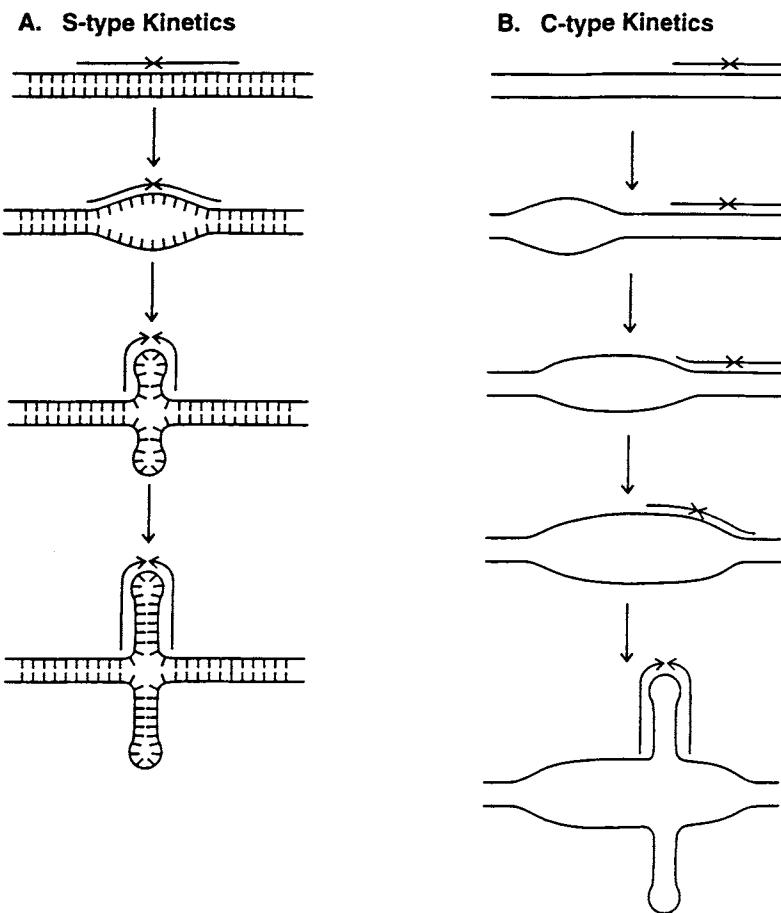


Figure 18. S-type and C-type mechanisms of cruciform formation. (A) For the S-type cruciform mechanism, 10 bp at the center of symmetry must unwind. Nucleation occurs with intrastrand hydrogen bond formation flanking the center of symmetry. Following nucleation, branch migration occurs in the extrusion process. (B) For C-type kinetics an A + T-rich region of DNA flanking the inverted repeat breathes, forming a denaturation bubble. The bubble becomes enlarged, encompassing the inverted repeat. Within this unwound region nucleation can occur, resulting in formation of a cruciform structure.

2. Effect of Base Composition on the Formation of Cruciforms

The rate of unwinding the center of an inverted repeat depends on its base composition as well as temperature, ionic strength, and superhelical density of the DNA. The thermal stability of the central 10 bp is important in defining the rate of

cruciform formation. In general, there is a correlation between the T_m of the entire inverted repeat and the rate of cruciform formation: DNAs with lower T_m values form cruciforms more easily than sequences with higher T_m . As expected, there is a cooperative relationship between superhelical density and temperature for cruciform formation.

3. C-Type Cruciform Formation

The characteristics of cruciform formation discussed above pertain to solutions of physiological ionic strength. This type of transition has been called the S-type (where S refers to salt-dependent). The S-type transition is dependent on supercoiling, temperature, ionic conditions, as well as divalent cation.^{126–128} A second mechanism, called the C-type, occurs in solutions lacking salt^{129,130} (Figure 18B). Under C-type conditions, the rate of cruciform formation is independent of base composition.

C-type behavior is due to A + T rich DNA sequences, called C-type inducing sequences, located within several hundred base pairs of the inverted repeat.^{130,131} Under conditions of low ionic strength, the A + T-rich C-type inducing sequence unwinds, forming a large unwound region that probably includes the inverted repeat.¹³² Nucleation leading to hairpin formation can occur within this unwound region (Figure 18B). A C-type transmitting sequence must be present between the C-type inducing sequence and the inverted repeat for C-type cruciform formation.¹³¹ The presence of a G₄C₄ block between the A + T-rich C-type inducing sequence and the inverted repeat will prevent C-type cruciform formation. Presumably, the high thermal stability of G + C-rich DNA resists melting and prevents unwinding through the inverted repeat.

4. The Removal of Cruciforms from Supercoiled DNA

There are three ways in which cruciforms can be converted back to the linear form: by introduction of positive DNA supercoils, heating above the T_m of the inverted repeat, and removal by transcription or DNA replication.

Introduction of positive supercoils. The introduction of positive supercoils into DNA will drive cruciforms back into the linear form, since they require negative supercoiling for stabilization. Positive supercoils can be introduced by the binding of ethidium bromide, which intercalates into DNA. Upon the removal of ethidium bromide, the DNA becomes reequilibrated with negative supercoils, and if ethidium bromide is removed at 0–4°C, this low temperature prevents the formation of cruciforms (except for (AT)_n inverted repeats).

Heating above the T_m . Cruciforms are stable in supercoiled DNA below the T_m of the hairpin arm. Upon incubation above the T_m , the hairpin arm is melted, and

upon cooling, the inverted repeat returns to the linear form, presumably because two nucleation events are provided by the plasmid DNA flanking the inverted repeat. Reannealing from the ends of the inverted repeat results in rapid formation of linear DNA. It is probably the kinetics of this process, compared to the slower rate of intramolecular base pairing, that drives the formation of linear DNA, despite the fact that cruciforms are thermodynamically stable in supercoiled DNA.

The removal of cruciforms by transcription and DNA replication. Transcription by RNA polymerase through a cruciform will reconvert an inverted repeat into the linear form. As RNA polymerase traverses the cruciform, it will necessarily unwind the base pairing in the hairpin arms, and the duplex DNA behind the RNA polymerase provides a nucleation site for hybridization of the complementary strands, as demonstrated experimentally by Morales et al.¹³³ Transcription and replication are probably responsible for removing cruciforms from DNA *in vivo*.¹³⁴

C. Cruciform Structure

Cruciforms have two structural characteristics, the four-way DNA junction and the stem loops.

1. Four-Way Junctions

The three-dimensional structure of four-way DNA junctions (Holliday structure) has been a field of intense investigation (Figure 19). Most of the structural analysis of four-way DNA junctions has been modeled on small stable junctions composed of four synthetic oligonucleotides (reviewed by Lilley and Clegg¹³⁵ and Wemmer et al.¹³⁶). Using these stable junctions, it was found that (1) four-way junctions are normally fully base paired¹³⁶; (2) all four branches are in the right-handed B-conformation^{137,138}; (3) there is twofold symmetry,¹³⁹ which is in agreement with the stacked X-conformation¹³⁸ (see below), and the most likely isomer is the antiparallel conformer.^{140,141} The structure depends critically on the DNA sequence, mainly at the junction, which determines the distribution of the stereoisomers,^{142,143} and on the type and amount of counterion used in the solutions, which determines the geometry of the helices.¹⁴⁴ In the absence of salt, the junction is in an extended conformation, probably planar with fourfold symmetry, with unstacked bases at the junction.¹⁴⁵ Micromoles of Mg²⁺ enable the four-way junction to adopt a more compact, X-shaped structure with twofold symmetry, with pairwise coaxial stacking of helices and apparently no unpaired bases.¹⁴⁵ In the presence of Na⁺ (≥ 50 mM), the structure is similar to that in the presence of Mg²⁺, in that it is also compact (X-shaped)^{135,146}; however, the structure has only imperfect twofold symmetry,¹⁴⁷ and the junction bases are still unstacked.^{144,145}

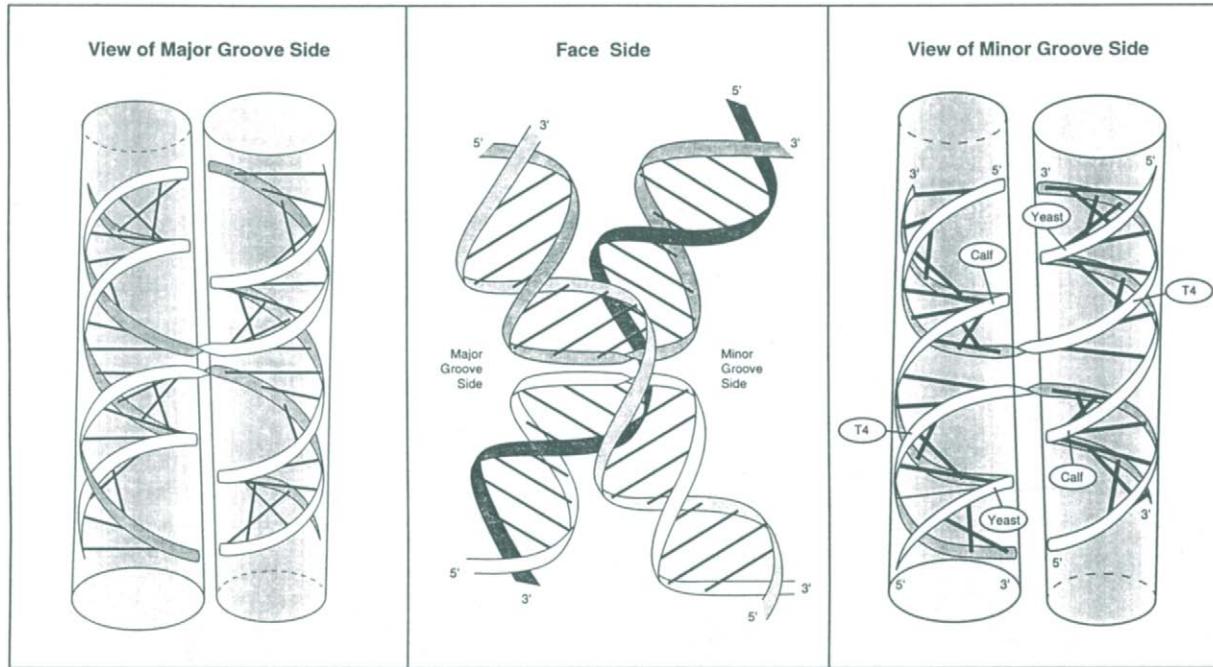


Figure 19. The Holliday or four-way junction. (**Center**) Helical representation of a Holliday junction. The four individual strands have different shadings. (**Left**) This representation shows the major groove side of the two DNA helices. In a Holliday junction, no Watson-Crick hydrogen bonds are broken, and all base pairs remain stacked. (**Right**) This representation shows the minor groove side of a Holliday junction. Also shown are the sites of cleavage of the yeast, calf, and T4 Holliday structure resolvases. (Figure modified with permission from Lilley and Clegg.)¹³⁵ Reproduced with permission from the *Annual Review of Biophysics and Biomolecular Structure*, Volume 22. ©1993 by Annual Reviews, Inc.

It has been reasoned that the more physiological structure is the more compact (X-shaped), rather than the extended conformation.^{135,146,148}

The structural model of four-way DNA junctions has recently had several interesting developments, namely that the given structure of a particular set of sequences does not have a stable steady-state structure.¹⁴⁹ It is interesting to note that cruciforms can induce DNA bending,¹⁵⁰ and that four-way junctions are relatively flexible.¹⁵¹ Regardless of the geometric conformation, each four-way DNA junction possesses both a minor groove face and a major groove face. On one face the four base pairs at the point of strand exchange all present minor groove edges, and on the other side, the corresponding major groove edges are presented.¹⁴⁵

2. *Stem Loops*

Single-stranded loops at the tip of cruciform arms or DNA hairpins are sensitive to single-strand nucleases. For direct inverted repeats, with no intervening sequence between the repeats, in the cruciform conformation the loops contain two or three unpaired bases.^{152,153} For nondirect inverted repeats, the loop size is dependent upon the length of the intervening sequence. Hairpin loops can involve base stacking,¹⁵⁴ non-Watson-Crick base pairing,^{155,156} and, possibly, extrahelical bases.¹⁵⁷ Hairpin loop conformation and dynamics are exquisitely sensitive to small changes in the bases in the loop and adjacent stem sequences.^{154,158} Similarly, the presence of hairpin loops can affect the stem structure¹⁵⁹ and may affect the flexure, writhe, and torsional strain experienced by the molecule. Intrahelical torsion may severely affect the conformation of a four-way junction.¹⁵¹ Hence, for cruciforms with short arms, the presence of a hairpin loop, as opposed to a free end, may affect the overall conformation of the four-way junction. The presence of loops could impart a certain stiffness to the DNA structure, which may be important for protein recognition of the target, such as a cruciform/hairpin as opposed to a Holliday junction, a DNA cross-over or a looped (wrapped) DNA molecule.¹⁶⁰

3. *Four-Way DNA Junctions, Holliday Junctions, and DNA Cross-Overs*

Holliday structures, believed to occur as recombination intermediates,¹⁶¹ contain a four-way DNA junction. In the case of the Holliday structure, it is a junction between two separate DNA molecules (whereas a cruciform is an intramolecular structure). DNA cross-overs are regions where two separate DNAs cross, as would occur at points where two double helices intersect by the process of supercoiling, looping, or folding, and at nucleosome linkers where DNA strands enter and exit the nucleosome.¹⁶² DNA cross-overs also resemble four-way DNA junctions, except that there is no covalent interaction between the two helices. Both Holliday

junctions and DNA cross-over structures, although similar, are biologically distinct from cruciforms.

4. Branch Migration of Four-Way DNA Junctions

Both the extrusion of cruciforms from inverted repeats as well as the formation of Holliday junctions during genetic recombination require branch migration. Branch migration involves the stepwise breakage and reformation of hydrogen bonds between base pairs as one strand is exchanged for another in the two arms of a four-way junction. The point of exchange occurs at the junction. This is a key step in most models of homologous recombination and is implicated in site-specific recombination. The number of hydrogen bonds formed equals the number broken, and the process can occur spontaneously. It occurs between regions of homology; in the case of cruciforms, branch migration is limited to the boundaries of the inverted repeat being extruded. Panyutin and Hsieh have shown that single base mismatches, such as those that occur in imperfect inverted repeats, can impede the process of branch migration.¹⁶³ Recent evidence indicates that the process of branch migration of DNA in four-way junctions, long thought to occur spontaneously at a rapid rate, occurs at a very slow rate.¹⁶⁴

The process of spontaneous branch migration has been modeled as a one-dimensional random walk in which there is an equal probability of moving forward or backward at each step. The rate of spontaneous branch migration is exceedingly sensitive to metal ions. The rate of branch migration increases dramatically at concentrations of MgCl₂ below 500 μM. In the absence of any metal ions, branch migration is exceedingly fast, with a step time of 50 μsec.¹⁶⁵ This increase in branch migration rate coincides with the loss of base stacking in the four-way junction over similar concentrations of the divalent ion. As described above, a constrained (immobile) four-way DNA junction adopts different conformations, depending on the metal ion present. In the absence of metal ions, the junction adopts a square-planar conformation. In the presence of multivalent ions, such as Mg²⁺, the junction assumes a stacked X structure, in which base stacking is retained through the cross-over point. It seems that the presence of multivalent metal ions at concentrations that induce base stacking at the junction also results in relatively slow rates of spontaneous branch migration.¹⁶⁵ This indicates that the structure of the Holliday junction is a critical determinant of the rate of spontaneous branch migration. The rate of branch migration in the presence of Na⁺ alone is similar to that observed in the presence of low concentrations (100 μM) of Mg²⁺, which probably reflects a loss of base stacking in both instances.

It has been proposed^{165,166} that the Holliday junction undergoes an isomerization at each migratory step, from a folded, stacked X structure to a square-planar conformation. This structural isomerization may involve the displacement of Mg²⁺ in the vicinity of the cross-over, leading to the disruption of base stacking at

the junction. Such conformational changes are likely to require energy, and therefore the process is slow.

Branch migration is known to affect binding, protein-induced DNA structural distortion, and endonucleolytic activity of phage, bacterial, and yeast Holliday junction resolvases.¹⁶⁷⁻¹⁷³ There are cellular proteins that can drive branch migration¹⁷⁴ (reviewed by Adams and West).¹⁷⁵

D. Assays for Cruciform Structures in DNA

Many assays for cruciforms have been developed. Many of these are very similar to assays applied to the analysis of Z-DNA, intramolecular triplex DNA structures, as well as other non-B-DNA structures. Table 3 summarizes the basics of these assays.

1. Application of the Psoralen Crosslinking Assay for Quantitating Cruciforms

Psoralen is an excellent probe for *in vivo* studies because it is freely permeable into eukaryotic cells and reasonably permeable into bacterial cells.¹⁸²⁻¹⁸⁴ Because of the low solubility of many psoralen derivatives and the relatively low binding constant to DNA, low levels of psoralen bind to DNA by intercalation in its "equilibrium dark-binding" mode, even when a solution is saturated with psoralen. On absorption of 360 nm light, covalent cyclobutane rings are formed between one or both ends of the psoralen molecule and pyrimidine bases. The extent of DNA cross-linking can be precisely controlled by varying either the light dose, the psoralen concentration, or both. Using the psoralen crosslinking assay, Zheng et al. demonstrated the superhelical density-dependent existence of a series of 106-bp cruciforms with A + T-rich centers in *E. coli*.¹³⁴

2. Nuclease Assays for Cruciforms in Cells

If a nuclease existed in cells that specifically recognized and cut cruciforms, then identification of specific cutting at inverted repeats should provide an indication of the existence of cruciforms. Panayotatos and Fontaine cloned the T7 endonuclease VII gene on a plasmid under transcriptional control of a regulated promoter.¹⁸⁶ When the gene was turned on and the nuclease was expressed in cells, cuts were made at the inverted repeats. In addition, cuts were made throughout the entire chromosome, resulting in complete destruction of the intracellular DNA. This experiment is certainly consistent with the interpretation that cruciforms existed in cells. It is difficult to rule out, however, that the nuclease was also recognizing other structural features in chromosomal DNA.

Table 3. Assays for Cruciform Structures

Assay	Description	Representative Reference
Single-strand nuclease cutting of cruciform loops	S1 nuclease is a single-stranded endonuclease that does not cut relaxed plasmid DNA containing the linear form of an inverted repeat. Digestion of supercoiled plasmid DNA, however, results in the introduction of a double strand break at the center of the inverted repeat. The location of the cut can be mapped by digestion with a restriction endonuclease, which cuts the plasmid DNA once at a known site. From analysis of the sizes of the two fragments on agarose gels, the S1 nuclease-sensitive site can be determined. The analysis of restriction-S1 fragments on DNA sequencing gels can identify the loop of a cruciform with base pair resolution.	Lilley ¹⁷⁶ Panayotatos and Wells ¹²³
Chemical modification at the loop of hairpin stems	The reactivity of the unpaired bases at the tips of hairpin arms to certain chemicals, including chloroacetaldehyde, osmium tetroxide, and diethylpyrocarbonate, provides an assay for cruciforms. Chemical cleavage (after modification) can be used to identify the site of modification with base pair resolution. Alternatively, sites of modification can be identified by primer extension procedures, since DNA polymerization will stop at the modified base. The products of either protocol are analyzed on a DNA sequencing gel.	Palecek ¹⁷⁷ Furlong et al. ¹⁷⁸
Analysis using cruciform resolvases	Resolvases, which cut cruciforms at the base of the arms, can be used to identify cruciforms in DNA.	Lilley and Kemper ¹⁷⁹
Psoralen interstrand cross-linking assay for cruciforms	Psoralen forms interstrand cross-links between the two strands of a DNA duplex on exposure to 360-nm light. The inverted repeat can be covalently locked into either the cruciform or linear form. Because the cross-linked hairpin arm is half the mass of the cross-linked linear form, these DNA species from cloned inverted repeats migrate very differently on a polyacrylamide gel, permitting analysis of the extent of cruciform formation. This assay is applicable <i>in vivo</i> . Cross-linking a cruciform also locks a bend, inherent in the four-way junction into DNA, that provides an additional assay for cruciforms.	Sinden et al. ¹⁸⁰ Zheng and Sinden ¹²⁰
DNA topological analysis	Cruciform formation results in the relaxation of one negative supercoil for about every 10.5 bp extruded into a cruciform. The change in electrophoretic mobility on an agarose gel provides a very sensitive assay for cruciform formation.	See Sinden ¹
Alteration of restriction cleavage	When cruciforms contain a restriction enzyme site at the center of the inverted repeat, digestion will occur in the linear but not the cruciform form.	Mizuuchi et al. ¹²⁴
Cruciform-binding antibodies	Cruciform-specific antibody binding, which can be detected by a variety of methods, provides a sensitive assay for cruciforms.	Frappier et al. ¹⁸¹

3. Detection of Cruciforms in Cells Using Chemical Probes

Certain chemicals that modify the unpaired bases at the loop of a hairpin arm can enter cells and preferentially modify these regions of DNA. Following modification, the DNA can be purified and analyzed to identify the site of modification. A disadvantage of this approach is that the cells are killed following treatment with many chemical probes. Osmium tetroxide (OsO_4), for example, severely damages living cells. This approach has been used by McClellan et al. to modify the loops of cruciform arms in a series of $(\text{TA})_n$ inverted repeats.¹⁸⁷ Sequences greater than 30 bp were chemically modified, but 24-bp sequences were not reactive. Moreover, the level of chemical modification was higher with increasing lengths of the inverted repeats. Since the level of negative supercoiling required for cruciform formation increases with shorter $(\text{TA})_n$ inverted repeats, the longest inverted repeats would be expected to exist as cruciforms at the highest levels in cells.

4. Analysis of DNA Supercoiling as an Indication of Cruciform Formation in Cells

The formation of a cruciform in a plasmid results in the relaxation of one negative supercoil for about every 10.5 bp of DNA extruded into a cruciform. *E. coli* maintains a precise level of negative supercoiling in cells. Therefore, the relaxation of negative supercoils in cells by the formation of a cruciform should theoretically induce a decrease in the linking number in plasmids containing extruded cruciforms and lead to the restoration of the original level of supercoiling by the action of DNA gyrase. Following plasmid purification from cells, providing that no change in the linking number occurred during purification, a population of topoisomers should be present that have more negative supercoils than plasmids in which the cruciform did not form. The increase in the number of supercoils upon cruciform formation *in vivo* (ΔLk_c) should be approximately equal to $\Delta Lk_c = N/10.5$, where N is the number of base pairs in the inverted repeat that have extruded into the cruciform. Typically, the formation of cruciforms *in vivo* does not occur in all topoisomers of a plasmid population. In the case where cruciforms formed in 50% of the topoisomers, rather than a single Gaussian distribution of topoisomers there would be two overlapping Gaussian distributions. The higher negatively supercoiled population, resulting from cruciform formation, would be shifted by ΔLk_c . Haniford and Puleyblank used this system to present one of the first indications of formation of cruciforms in *E. coli*.¹⁸⁸

5. Application of Cruciform Antibodies to the Analysis of Cruciforms in Eukaryotic Cells

Monoclonal antibodies have been raised and isolated that recognize structural features of cruciforms, but not the DNA sequence of specific inverted repeats. *In*

vitro antibody binding to DNA can be detected by the DNA filter binding or the retardation of cruciform-containing DNA during gel electrophoresis.¹⁸¹ When antibodies are applied to isolated nuclei, they can be detected by binding a second fluorescent antibody to the anticruciform antibody. The localization of the fluorescence by microscopy reveals the sites and relative abundance of cruciforms.

In eukaryotic cells, cruciforms are detected most strongly at the G₁/S boundary of the cell cycle, just before the beginning of the period of DNA synthesis (S phase). Ward et al. estimated that there might be as many as 3×10^5 cruciforms per eukaryotic nucleus and suggested that inverted repeats may form cruciforms as a prerequisite for the initiation of DNA replication.^{189,190} As discussed below, inverted repeats are quite common at origins of DNA replication. Perhaps the formation of a cruciform triggers initiation of DNA replication at a specific origin. Alternatively, cruciforms may simply accumulate at the G₁/S boundary, during the period when no replication is occurring. As soon as synthesis begins, cruciforms may be converted by replication through the inverted repeat back into the linear form.

Caution should be exercised in interpreting the cruciform antibody (and Z-DNA antibody, see Section V) studies of alternative conformations in eukaryotic cells. These analyses require the isolation and purification of nuclei to allow the antibody to enter the nucleus. The cells must be gently lysed, removing the nuclei from their natural environment. It has been shown by Cook et al. that even the gentlest isolation procedures result in significant changes in the chromatin organization of DNA.¹⁹¹ In eukaryotic cells, negative supercoiling in the bulk of chromosomal DNA is restrained by the organization of DNA into nucleosomes. Relaxed DNA should not support the formation of cruciforms (or Z-DNA or triplex structures). However, if the chromatin structure is disrupted and nucleosomes are lost, unrestrained supercoils may be introduced into DNA as an artifact and may not reflect the natural *in vivo* situation of the DNA. It is interesting to note that in addition to these studies on cruciforms, there are several reports of other non-B-DNA structures that appear to be dynamically regulated throughout the eukaryotic cell cycle, including triplex DNA,¹⁹² Z-DNA,^{193,194} and single-stranded DNA.^{195–198}

E. The Biology of Inverted Repeats

1. Protein Binding Sites

Many short inverted repeats (4–20 bp) represent the binding site for specific proteins. Restriction endonuclease cleavage/modification sites range from 4 to 10 bp in length. However, recognition and cleavage of these sites occur in the linear form, since most restriction sites are too short to form a cruciform (which requires a 3–4-bp loop at the end of the hairpin stem). Many DNA operator sequences have quasipalindromic symmetry, in which the inverted repeat is not perfectly symmet-

rical. The binding of dimeric repressor proteins usually involves recognition of a 6–10-bp DNA sequence by each monomer repressor.

2. Inverted Repeats and Other Sequence Elements Associated with Eukaryotic Replication Origins: Regulation of Replication Initiation

A pressing question in cellular and molecular biology is how the cell limits DNA synthesis to one round per cell cycle. The mechanism that inhibits reinitiation is not known, although chromatin conformation, chromatid pairing, and nuclear membrane permeability may be involved in the process. Replication of genomic DNA is restricted to the S phase of the cell cycle, and control mechanisms ensure that the entire genome is replicated only once per cell cycle.

Replication origins and replicators, the specific sequences that control the initiation of DNA replication, are poorly defined in mammalian cells.^{199,200} Unlike the small simple genomes of prokaryotes and viruses, the multichromosome genome of a mammalian cell is more complex. With this increased complexity one might expect increasingly complex or more numerous modes or levels of initiation regulation. However, despite the lack of a particular consensus sequence, there are certain types of sequences that are common to many replication origins of parasitic, prokaryotic, eukaryotic, and mammalian organisms, some of which are described below.

A- + T-rich sequences. The helical axis of DNA can be curved unidirectionally if there exists a series of A tracts that are phased in their spacing (reviewed by Hagerman²⁰¹; also see Section II above). Curved DNA is known to occur at prokaryotic, yeast, viral, and mitochondrial replication origins. Curved DNA has been associated with the DHFR and c-myc-associated origin regions, as well as monkey and human origin enriched sequences (*ors*).

DNA unwinding elements. Common to most origins are sequence elements that facilitate the unwinding of the DNA. DNA unwinding elements (DUEs) are specific but not unique sequences. They are usually A + T-rich, and their function is dependent upon the base stacking interactions.^{202,203} DUEs were first recognized as regions of single-strand nuclease hypersensitivity contained within the *E. coli*²⁰⁴ and yeast^{202,205} replication origins. Mutations in the DUE of the yeast autonomously replicating sequence (ARS) that inhibit DNA unwinding *in vitro* also inhibit ARS activity *in vivo*.²⁰²

DUEs are believed to function by unwinding the two strands to permit the entrance of the replication machinery to initiate replication, some of these regions extend to lengths up to 100 nucleotides. Recently the DUE of SV40 was found to be within the G + C-rich early region adjacent to the central inverted repeat.²⁰⁶ This corresponds directly to the site of initial primer synthesis.^{207,208} Cruciforms

may play a role in DNA melting, for the easily melted regions of pBR322 contain inverted repeats.^{209,210}

The Epstein-Barr virus replication origin has its DUE located at the origin of replication, oriP.²¹¹ Recent evidence indicates that both the dyad symmetry element and the family of inverted repeats are sensitive to single-strand nucleases.²¹¹ In duplex DNA the structure of the dyad symmetry element is a large single-stranded bubble containing a stem loop formed by the 65-bp dyad, whereas the family of repeats are in the cruciform conformation. Williams and Kowalski concluded that the intrinsic ability of the oriP elements to form alternative structures may be important in the initiation process, specifically facilitating the access of the replication machinery to the parental DNA strands.²¹¹ The unwound single-strand bubble containing a base paired hairpin is reminiscent of the single-strand initiation (*ssi*) signals of plasmids and phages. It is tempting to speculate that the unwound oriP is recognized by replication factors in a fashion similar to that of the *ssi* signals.

Inverted repeats. Inverted repeats are common sequence elements in many prokaryotic and eukaryotic replication origins (reviewed by Pearson et al.)¹¹⁸ and are important for the initiation of DNA replication in phages, plasmids, prokaryotes, and viruses of both prokaryotes and eukaryotes.

Hairpins in single-stranded bacteriophage genomes. Inverted repeats in single-strand DNA can form hairpins, and in single-stranded chromosomes these are important for replication. A region containing three inverted repeats of 44, 21, and 20 nt is required for the initiation of replication in bacteriophage G4 (see Table 4.2 in Sinden).¹ The hairpin structures of the 44-bp and 21-bp inverted repeats are required for this region to function as the origin. Following infection, the DnaG primase protein binds to the hairpin formed within the 20-bp inverted repeat as a prerequisite for initiation of synthesis of an RNA primer.

Eukaryotic viruses. Inverted repeats are associated with many eukaryotic viral origins of replication. For example, the SV40 origin of replication contains a perfect 27-bp inverted repeat and an imperfect 15-bp “early palindrome.” The linear duplex forms of the inverted repeats are required for replication. Herpes simplex virus contains *ori*_{L1}, a 144-bp A + T-rich inverted repeat with a perfectly symmetrical central 20-bp region,²¹² and *ori*_{L2}, a 136-bp inverted repeat that is similar to *ori*_{L1}.²¹³

Plasmid pT181: Is a cruciform cut to begin replication? Plasmid pT181 contains a small inverted repeat that constitutes the origin of replication. This inverted repeat can form a cruciform *in vivo*, and it may be involved in the initiation of replication.²¹⁴ RepC, an initiation protein, binds to the origin of single-stranded or double-stranded DNA and introduces a nick at the center of the inverted repeat.

Replication begins at the nick. The binding of the RepC protein is enhanced by the formation of an unwound structure, possibly the cruciform, at the origin.

Inverted repeats in eukaryotic genomes. To investigate the sequence/structure requirements of higher eukaryotic replication origins, numerous laboratories isolated libraries of early replicating sequences that are enriched in replication origins which are activated at the onset of the S phase. The nascent fragments, ranging from several hundred base pairs to over 1 kb, were cloned, generating a library of early replicating sequences. These sequences, by the nature of their isolation, should contain replication origins at or near their center and thus are termed "origin-enriched sequences. These were isolated from avian, mouse, monkey, and human cells (reviewed by Pearson et al.).¹¹⁸ Some of these cloned sequences were capable of autonomous replication upon transfection into mammalian cells. Sequence analysis did not reveal a single major consensus sequence among the clones or between the libraries. However, in each of the above-mentioned libraries the sequences were enriched with both short inverted repeats as well as A + T-rich tracts. In addition, the origin of bidirectional replication of the dihydrotoluate reductase (*DHFR*) gene and the *c-myc*-associated replication origin are both known to contain inverted repeats (reviewed by Pearson et al.).¹¹⁸

3. Mammalian Cells: The Effect of Anti-Cruciform-DNA Antibodies

Monoclonal antibodies have been produced with unique specificity to cruciform DNA structures.^{181,215} These antibodies recognize conformational determinants specific to DNA cruciforms and do not bind linear double-stranded DNA, linear single-stranded DNA containing a stem loop structure or tRNA. The binding site of these antibodies has been mapped to the four-way (elbow-like) junction at the base of the cruciform.^{215,216}

Introduction of the anticruciform DNA antibodies into a permeabilized cell system, which is capable of carrying out DNA replication, resulted in a 2- to 11-fold enhancement of DNA synthesis.²¹⁷ An enhanced replication of known early replicating sequences such as *ors8*, *DHFR*, and *c-myc* was also detected.²¹⁷ Taken with the above mentioned precaution of *in vivo* analysis (see above), this effect was apparently caused by the antibody supposedly stabilizing cruciforms encountered near the origins of replication and allowing multiple initiations to occur at these sites.

Using the anti-cruciform DNA antibodies in the same system, Zannis-Hadjopoulos and colleagues^{189,190} were able to quantify the number of cruciform structures in living cells by fluorescent flow cytometry. Two major populations of cruciforms were observed throughout the cell cycle, each with an estimated 0.6×10^5 and 3×10^5 cruciforms per cell, respectively. Cruciforms were observed in a bimodal distribution throughout the S phase, their numbers reaching a maximum at the G₁/S boundary. The second wave occurred at 4 hours into S phase, but at a

lower level than that observed at the G₁/S boundary. The timing of these waves was coincident with both the maximum rate of DNA synthesis and the relative enhancement of DNA synthesis by the antibody.²¹⁷ A limited number of cruciforms were detected in G₂/M nuclei. These data suggest that the formation of cruciforms is cell-cycle regulated. It also suggests that both cruciforms and active replication origins are grouped closely together in discrete regions within the nucleus in early S phase. The above observations support the hypothesis that certain inverted repeats may represent potential initiation sites for DNA replication.

4. Inverted Repeats and Gene Amplification

Very long inverted repeats in mammalian cells have been associated with gene amplification of oncogenes (e.g., N-myc, c-myc, erb-B, sis, ras) and of genes involved in drug resistance, such as the adenylate deaminase (AMPD), adenine phosphoribosyltransferase (APRT), carbamoyl-phosphate synthetase/aspartate carbamoyl transferase/dihydroorotase (CAD), and dihydrofolate reductase (DHFR) genes (reviewed by Fried et al.²¹⁸ and Windle and Wahl²¹⁹). Treatment of mammalian cells with replication inhibitors and certain drugs can induce gene amplification, giving rise to inverted and tandem repetitions of genes and their surrounding regions. Such genetic amplification allows for multiple copies (>1000 in the case of DHFR) and overexpression of the gene product. The palindromic arms of the amplified mammalian inverted duplications are extremely large, in excess of 100 kb. In each case studied, the arms are separated by a stretch of noninverted DNA of approximately 150–1000 bp.²¹⁸ The formation of the large inverted repeats has been explained by way of an extra-chromosomal double rolling circle model.^{169,218} This model requires that an origin of replication be in the vicinity of the DNA region to be amplified. Interestingly, sequence analysis of the sites of the recombination joints of inverted duplications in amplified DNA reveals that the DNA contained A + T-rich sequences as well as short inverted repeats with the potential to form hairpin (cruciform) structures, both of which are common to origins.^{169,220,221}

The long inverted repeats of amplified DNA do not form cruciforms *in vivo*.²²¹ However, the long inverted repeats do form cruciforms *in vitro* following DNA isolation.²²¹ Processes of genomic DNA isolation can induce topological perturbations in the DNA²²². The amplification and synthesis of long (plasmid-length) inverted repeats can be mimicked in an *in vitro* T antigen/SV40 origin-dependent replication/amplification assay by using extracts from carcinogen-treated HeLa cells.²²³⁻²²⁵

5. Inverted Repeats, Cruciforms, Z-DNA, and Nucleosomes

Nucleosomes interfere with the binding of initiation factors to promoters²²⁶ and origins of replication.^{227,228} A yeast ARS placed within the nucleosome has

severely reduced replication activity compared to its normal location in the linker region.²²⁸ Histones bind poorly to inverted repeats^{229,230} or cruciform DNA structures,²³¹⁻²³³ and it is likely that cruciform structures exist in the spacer region between nucleosomes.²³¹ Cruciforms may play a role in nucleosome phasing, such that they expose nucleosome-free DNA sequences, making them accessible to DNA binding proteins specific for transcription, recombination, and/or replication.²³¹⁻²³⁴

The formation of Z-DNA may also play a similar role in nucleosome phasing.²³⁵⁻²³⁹ The SV40 viral replication origin region, in addition to the central inverted repeat and adjacent AT tract, contains 21-bp and 72-bp repeats. Both the 21-bp and the 72-bp repeats enhance the replication efficiency, in an orientation-independent manner. The position of the repeats is important.^{240,241} The 21-bp and the 72-bp repeats are believed to enhance replication by maintaining a nucleosome-free region at the SV40 origin.²⁴²⁻²⁴⁴ Interestingly both the 21-bp and 72-bp repeat regions have been reported to adopt the Z-DNA conformation.²⁴⁵⁻²⁴⁹

6. *Inverted Repeats, Stem Loops, Cruciforms, and Transcription*

Inverted repeats are present in a number of regulatory regions of genes, ranging from simple plasmid borne antibiotic resistance genes²⁵⁰ and viral genes²⁵¹ to well-studied eukaryotic genes.²⁵²⁻²⁵⁵ DNA sequences that can potentially form secondary structures are frequently localized to promoter regions, suggesting their role in gene regulation^{250,256,257} (reviewed by Horwitz and Loeb).²⁵⁸

Involvement of an inverted repeat in N4 virion RNA polymerase promoter recognition. The bacteriophage N4 RNA polymerase, which is packaged into the virion, is required for early transcription. Early transcription requires a supercoiled template, on which the *E. coli* SSB protein acts as a transcriptional activator.²⁵⁹ The promoters of these early genes contain inverted repeats necessary for transcription. Remarkably, the inverted repeat symmetry and not a defined base sequence is required for transcription.²⁵⁷ A model has been proposed in which DNA is supercoiled by DNA gyrase, driving the formation of a short cruciform structure stabilized by a particularly stable hairpin loop in one strand, containing the sequence CGAAG. The SSB protein removes any potential alternative secondary structure in the strand containing the sequence CTTCG. The N4 RNA polymerase then binds to the strand containing the hairpin stem.²⁵⁷ This system represents probably the most convincing evidence for a system in which a cruciform structure is involved in the regulation of gene expression.

Possible involvement of cruciforms in transcription regulation. The twin-domain model of transcriptional supercoiling¹¹³ predicts the formation of positive supercoils in front of the transcription complex, and the generation of negative supercoils behind it.^{260,261} Supercoil waves generated by transcription may cause

the reversal of negative supercoil-induced altered structures to the B-form. Based on the proposition that supercoiling can be affected by protein tracking mechanisms such as transcription, several researchers investigated the effect of cruciforms on such processes. There have been several reports of cruciforms regulating DNA transcription: examples of cruciform extrusion driven by transcription²⁶²; cruciform absorption by transcription¹³³; and transcription inhibition by the presence of cruciforms.^{263–265} There is a report of a cruciform-inhibiting transcription and the relief of inhibition by the addition of HMG1.¹⁷³ HMG1 binds to DNA cruciforms²⁶⁶ (and see below). Cruciforms on negatively supercoiled templates stalled the progress of RNA polymerases. The addition of HMG1 to the reaction relieved the block by binding to the cruciform and altering the DNA conformation to permit the progression of the RNA polymerase. Apparently HMG1 was able to bind and force the reabsorption of the extruded cruciform to the linear form inverted repeat, thus allowing the passage of the transcription complex. HMG1 can also remove the transcriptional block caused by left-handed Z-DNA on a supercoiled template.²⁶⁷

ADP-ribosyltransferase, implicated in DNA repair, binds specifically to DNA four-way junctions.²⁶⁸ It has been proposed that this structure-specific mode of binding by the enzyme may be linked to its autoregulation of expression.²⁶⁹ The promoter structure of the ADP-ribosyltransferase gene contains several noteworthy inverted repeats, which could form cruciform structures. The functional significance of these inverted repeats was revealed by deletion analysis; removal of one inverted repeat resulted in diminished promoter function, whereas removal of the other increased promoter activity. Overexpression by a heterologous promoter of ADP-ribosyltransferase led to repression of the endogenous promoter, suggesting that the ADP-ribosyltransferase gene product acts as a negative modulator of its own promoter.

7. Other Biological Phenomena Associated with Cruciforms

Cruciform extrusion and regulation of superhelicity and protein–DNA interaction. The extent of supercoiling is known to affect transcription, recombination, and replication.^{261,270} Since cruciform extrusion causes an effective relaxation in DNA distal to it (the number of helical turns relaxed is approximately equal to the number of helical turns contained in the stem-loop of the cruciform)²⁷¹ the process of cruciform formation may be a mechanism regulating local superhelicity. In this manner, cruciform extrusion or melting may regulate the recognition and binding of proteins to specific sequences proximal to the cruciform.²⁶⁴ A particular instance of this was reported by Pearson et al.²⁷² It appears that the presence of a cruciform stabilizes the sequence-specific binding of a protein(s) at a site proximal to it. Presumably the cruciform altered the structure of the target sequence such that it was more efficiently recognized and/or bound by the protein. Cruciforms are known to affect structural alterations in the flanking sequences.^{150,273–275} This

mode of transmission along the DNA molecule has previously been observed (refs. 130 and 275, and references therein) and is referred to as "telestability."²⁷⁶ This transmission has been demonstrated to occur over long distances and can affect protein-DNA interactions such as the interaction of RNA polymerase with promoters,²⁷⁷ and nuclease specificity.⁵⁵ Horwitz found that the resulting decrease in superhelicity by cruciform extrusion at one plasmid based promoter is below the optimum for expression from another promoter on the same plasmid, indicating that cruciform formation can act at a distance.²⁶⁴

Cruciform-specific proteins. It has long been suggested that cruciform structures may form transiently under physiological conditions and serve as the recognition site for initiator or other protein factors²⁷⁸ (reviewed by Pearson et al.).¹¹⁸ Presented below is a discussion of proteins that recognize and bind DNA in a structure-specific fashion and are likely to be involved in transcription or replication.

Prokaryotic four-way junction resolvases. There are several proteins with nuclease activity that interact with four-way junctions. Resolvases are ubiquitous enzymes that cut the four-way junction.^{124,279-284} These enzymes cut 4–5 bp up the stem on either the 3' or 5' end of a cruciform arm (see Figure 19). Cuts are always made on opposite sides of the four-way junction to resolve the recombinant into two symmetrical molecules. Several of these have been analyzed in great detail by the use of hydroxyl radical cleavage of the DNA complexed with the proteins: the bacteriophage encoded resolvases T4 endonuclease VII²⁸⁵ and T7 endonuclease I,²⁸⁶ and the *E. coli* RuvC resolvase.¹⁷² The T7 endonuclease I and the T4 endonuclease VII contact the DNA backbone without detectably altering its structure. On the other hand, by hydroxyl radical analysis, the RuvC resolvase¹⁷² reveals only structural alterations, apparent as increases in radical cleavage intensity, but no protection footprint. Since a sequence dependence of the cleavage reaction for the RuvC resolvase was demonstrated, this protein might not interact directly with the DNA backbone.¹⁷² In the hydroxyl radical cleavage patterns of the bacteriophage endonuclease–cruciform complexes, only two diametrically opposed (T4) or all four junction strands (T7) are protected. T4, yeast, and calf thymus resolving enzymes interact with the minor groove face of junctions.²⁸⁷

Mammalian cruciform-DNA-binding proteins: HMG1, HMG2, and the HMG box. A number of mammalian cellular proteins have been reported that recognize the Holliday-like four-way junctions of DNA (reviewed by Duckett et al.)²⁸⁸; among them is the ubiquitous HMG1.²⁸⁹ HMG1 is an abundant nuclear protein with $>10^6$ molecules of HMG1 per cell.²⁹⁰⁻²⁹² HMG1 has been demonstrated to bind to gene regulatory regions, such as promoters, and replication origins with greater affinity than to "junk DNA."²⁹³ The high mobility group protein, HMG1, binds to four-way junctions.²⁸⁹ It has been reported that HMG1 can protect the single-stranded tips of cruciform stems from S1 nuclease digestion,¹⁷³ but

not from T4 endonuclease VII, which interacts at the junction.²⁸⁷ Not only is HMG1 able to bind to specific DNA structures, it is also capable of distorting the conformation upon binding. It has been reported to form beaded structures^{293,294} as well as unwind supercoiled DNAs.²⁹⁵⁻²⁹⁸ Recently several groups reported the ability of HMG1 to bend DNA such that there is facilitated T4 ligase-mediated circularization.^{299,300} Such circularization assays have been carried out for several proteins, including the bacterial HU protein.³⁰¹ Several authors have suggested that although HMG1 may not be similar to the bacterial HU protein at the sequence level, it certainly may be the mammalian functional equivalent.^{299,302} It has been proposed that the cellular function of HMG1 and HMG2 is to bind to folded regions of DNA, and thus aid in chromatin compaction.^{293,294,303} HMG1 and HMG2 proteins have been implicated in both transcription^{173,267,304-307} and replication processes.^{290,308-310}

A protein-DNA binding motif known as the HMG box,^{78,311,312} which is present in many different proteins, binds different forms of DNA, such as single-stranded DNA,³¹³ sharp angles in DNA,³¹⁴ and four-way DNA junctions.^{289,315,316} Clearly, a wide variety of altered DNA structures can be recognized and bound by HMG box proteins. In addition to binding to four-way DNA junctions, HMG1 binds to B-Z DNA junctions.²⁶⁶ Both posttranscriptionally modified HMG1 and 2 were isolated as Z-DNA-binding proteins. By using a brominated nucleic acid as a probe,³¹⁷ it was later discovered that these proteins were preferentially binding to brominated Z-DNA and not to nonbrominated DNA.^{318,319} Recently Lippard and co-workers isolated HMG1¹⁵¹ as well as other HMG box proteins that specifically recognize cisplatin DNA adducts.³²⁰

The HMG box is a domain of positively charged protein sequences characterized by one of the internal repeats contained in HMG1. The HMG1 protein is composed of three regions, consisting of two internal homologous repeats (HMG boxes) and an acidic tail.³²¹ The internal repeats consist of about 80 amino acids that are rich in basic and aromatic residues. The acidic tail is composed almost entirely of about 30 aspartates and glutamates. Although there is no apparent absolute conservation of the sequence for the HMG boxes, it is likely that they all have a similar structure. The HMG box is necessary and sufficient to bind DNA.³²¹ Both the whole HMG1 and its individual HMG boxes bind in a structure-specific fashion that is sequence independent. All proteins that contain HMG boxes bind to DNA. Some HMG box proteins recognize specific sequences that are generally A + T-rich, whereas others are indifferent to sequence (see discussion below).

The family of HMG box proteins includes proteins that may play key roles in transcription,^{322,323} replication,^{171,324} DNA repair,^{151,318,319} V-D-J immunoglobulin gene recombination,³²⁵ and sex determination (SRY).^{314,326} That the HMG box, a single motif, can be used for so many varied roles indicates a new class of nucleic acid binding proteins. Of the many proteins that contain HMG boxes, there are several classes: those that contain one HMG box, such as SRY³¹⁵; those that contain two, such as HMG1; and the human upstream binding factor

(hUBF)³²² and the *Xenopus* UBF,³²⁷ which contain four and five HMG boxes, respectively. The pressing questions are: (1) What is it that gives some HMG box proteins the ability to recognize in a site-specific manner? and (2) Is this recognition mediated by the rest of the polypeptide? A general trend is emerging—proteins that contain a single HMG box seem to bind with strong sequence specificity, and those with multiple HMG boxes bind in a sequence-tolerant fashion.³²⁸ The HMG1 protein itself contains two HMG boxes; this class of proteins binds solely in a structure-specific fashion. The human and *Xenopus* UBF proteins bind with little sequence specificity, and the *Xenopus* UBF, unlike other HMG box proteins, can bind to tRNA as well as DNA four-way junctions.³²⁷ The one-box class binds with significant sequence specificity as well as structure specificity. The sex-determining factor, SRY, contains one HMG box and binds specifically to DNAs containing its target sequence, AACAAAG. However, when SRY binds to its target it produces a sharp bend in the DNA; thus the interaction does contain some structure specificity.³¹⁴

The abundance of HMG1 and 2 indicates that they must be playing a role that is required continuously (such as chromatin structure), or that there are several roles within the cell, and these roles are mediated by the variety of modifications possible. More HMG box proteins will likely be discovered, and further research into the effect of modifications will be forthcoming.

Other four-way junction binding proteins. One more cruciform-specific binding protein (CBP) has been analyzed by the hydroxyl radical technique. CBP from human cells recognizes four-way DNA junctions in a structure-specific fashion²⁷² and is void of nuclease activity. It is apparent from the footprinting pattern of CBP¹⁶⁰ that the mode of interaction differs significantly from that of the bacteriophage encoded resolvases (see above) T4 endonuclease VII,²⁸⁵ and T7 endonuclease I,²⁸⁶ and the *E. coli* RuvC resolvase¹⁷²: (1) CBP interacts with the junction, giving clear areas of protection and simultaneously introducing several changes in the fine structure of the cruciform. (2) Whereas the hydroxyl radical cleavage patterns of the bacteriophage endonuclease–cruciform complexes show protection on only two diametrically opposed (T4) or all four junction strands (T7), the hydroxyl radical cleavage patterns of the CBP–DNA complex reveal protection on both junction strands of stem loops as well as both strands of the branch arms.^{285,286} This indicates that CBP apparently interacts at both major and minor groove faces of the cruciform junction. In addition, the fact that CBP protects both junction strands as well as both strands of branch arms indicates that CBP apparently interacts at both major and minor groove faces of the cruciform junctions; in contrast, T4, yeast, and calf thymus resolving enzymes interact with the minor groove face of junctions.²⁸⁷ Apparently CBP provides a novel type of cruciform DNA–protein interaction, in that there are firm contacts with the sugar phosphate backbone (protection) as well as structural alterations of the cruciform substrate, both reflecting a putative cruciform-stabilizing function of CBP in the cell. The

ability to structurally alter the DNA by binding of CBP provides a putative role for it in preparation of DNA for the processes of replication, transcription, or recombination. The asymmetric binding further predicts that there may well be a specific orientation of CBP required for an interaction with other proteins at a functional DNA element.

Several proteins have been implicated in binding specifically to DNA cross-overs, which occur at points of DNA looping and folding, and at nucleosome linkers.¹⁶² The histone H1, like HMG1, can also bind specifically to DNA cross-overs.^{328,329} Interestingly, the proteins are known to associate with each other,^{330,331} and it is known that HMG1 modulates the histone H1-induced condensation of DNA³³² and facilitates nucleosome assembly.³³³

Several reports indicate that eukaryotic DNA topoisomerases may specifically recognize DNA cruciforms.^{334,335} The vaccinia virus topoisomerase I has been shown to resolve Holliday junctions.³³⁶ This enzyme has a specificity for the site CCCTT; four-way junctions with these sites in two juxtaposed arms are recognized and cut by the enzyme. The cutting is essentially identical to that of type I topoisomerases in that it takes place by concerted transesterifications at the two CCCTT sites rather than by hydrolysis.

Topoisomerase II actually cleaves DNA hairpins,³³⁷ such that the cut site is one nucleotide 3' of the hairpin base. Recognition of the hairpin was sequence independent and required a double-stranded/single-stranded junction. Base pairing up to the base of the hairpin on the 5' side had no effect on the scission of the hairpin, whereas base pairing on the 3' end inhibited cutting. The requirement for a single stranded region was delimited to two to four nucleotides on the 3' side of the hairpin. Topoisomerase II has recently been demonstrated to specifically recognize and bind to Z-DNA³³⁸; the relationship between these different binding capabilities is not known. It may be that topoisomerase recognizes such structures as a means for regulating supercoil tension during processes of transcription, replication repair, or recombination.

Hairpin and stem loop binding proteins. In the human enkephalin gene enhancer there is an imperfect palindrome of 23 bp³³⁹ that contains both cAMP-responsive elements (CRE-1 and CRE-2). *In vitro* studies reveal that short oligonucleotides of this enhancer are able to undergo reversible conformational transitions from duplex to the hairpin structures.³³⁹ Although CRE-binding protein binds weakly to the duplex palindrome, it binds specifically and with high affinity to the G-T mispair in the hairpin formed by one of the two strands.³⁴⁰ Interestingly, the hairpin contains an alternative CRE binding site, which contains two G-T mispairings, in contrast to the A-T and G-C pairs found in the duplex site. Furthermore, a T in the mispaired positions is required for the high affinity recognition, as a fully base paired hairpin is not recognized. Thus the binding of CRE-binding protein is both sequence and structure specific.

There are proteins that bind specifically to single-stranded DNA that lies between two members of an inverted repeat.^{341,342} Such a situation occurs at the early palindrome of the SV40 virus, as there are two different human cellular proteins that bind to this palindrome.³⁴¹ Each binds with sequence specificity to one of the single palindromic strands. One was discovered to bind in a cell-cycle-regulated fashion. The other protein, binding to the opposite strand with some sequence specificity, is the replication protein A, a single-strand-binding protein.^{341,343} It has been proposed that the particular conformation of the loop structure, which is unique to each arm of a particular cruciform, can be specifically recognized by proteins.

Since each loop conformation is distinct, specific recognition of a particular loop would be determined by the loop sequence and perhaps the stem sequence (i.e., the environmental context of the loops). Alternatively, it may be the specific mismatch base pairings that occur in the stem loop(s) of a particular imperfect inverted repeat that are the recognition structures of proteins.³⁴⁰

The ubiquitous nature of short inverted repeats in the genomes of eukaryotic organisms suggests that these sequences were evolutionally maintained. The multitude of proteins that can interact with both sequence and structure specificity with cruciform DNAs suggest that cruciforms play biological roles in the cell.

V. LEFT-HANDED Z-DNA

A. Introduction

In 1970 Mitsui et al. suggested that the polymer poly d(I-C)-poly d(I-C) could adopt a left-handed conformation under specific conditions, based on the X-ray diffraction pattern and circular dichroism (CD) spectra of poly d(I-C)-poly d(I-C).³⁴⁴ However, much later Sutherland and Griffin³⁴⁵ found that the polymer analyzed by Mitsui et al. adopted an unusual non-B-DNA right-handed configuration called D-DNA. In 1972, Pohl and Jovin presented the CD spectrum of a left-handed alternating copolymer poly d(G-C)-poly d(G-C).³⁴⁶ This CD spectrum was very different from that for classical B-form DNA. The classical, characteristic B-form CD DNA spectrum shows a positive ellipticity (Θ) peak at about 270–280 nm, a negative ellipticity at about 250 nm, and a cross-over point (where the ellipticity equals zero) at about 260 nm. The Z-DNA spectrum shows an inversion of the peaks relative to the spectrum of B-DNA. In 1979 Rich and co-workers solved the X-ray crystal structure of d(CpGpCpGpCpG).³⁴⁷ Quite unexpectedly, this 6-bp oligonucleotide (the first crystal structure of a DNA molecule to be solved) existed as a left-handed helix.

Following the realization that certain alternating purine pyrimidine sequences could exist as a left-handed helix, an enormous scientific effort ensued to characterize the DNA sequences and environmental conditions required for Z-DNA for-

mation. The discovery of Z-DNA is important because the ability of DNA to adopt non-Watson-Crick structures could have profound implications for the processes of replication, recombination, or transcription. Z-DNA can exist readily in bacterial cells, and although sequences that can form Z-DNA are widely found in human DNA, nature has yet to divulge the biological role of Z-DNA in eukaryotic cells.

B. The Structure of Z-DNA

1. The Left-Handed Helix

There are only a few similarities between B-DNA and Z-DNA, including the double-stranded structure, antiparallel orientation of the two strands, and Watson-Crick hydrogen bonding.³⁴⁸ Apart from these similarities, there are a number of differences between B-DNA and Z-DNA. First, there is a zigzag phosphate backbone in Z-DNA compared to a smooth backbone in B-DNA. Second, B-DNA has distinct major and minor grooves. In Z-DNA, the major groove has all but disappeared into a nearly flat surface. The one visible groove is deep and narrow. This groove is structurally analogous to the minor groove in B-DNA. Third, the Z-DNA helix, at an 18-Å diameter, is narrower than B-DNA (20 Å in diameter). Fourth, the helix repeat in Z-DNA is 12 bp per turn, compared to 10.5 per turn for right-handed B-DNA. Finally, in B-form DNA the helix pitch is, on average, 36 Å, with an average rise of 3.4 Å. The helix pitch of Z-DNA is 44.6 Å, with a rise of 3.72 Å. (The helix parameters for B-DNA and Z-DNA are listed in Table 2.)

Two additional major structural differences between B-DNA and Z-DNA are the sugar pucker and the configuration of the glycosidic bond in deoxyguanosine. In B-form DNA all sugar residues exist in the C2' endo configuration. In Z-DNA the sugar pucker for dC remains C2' endo, but the pucker changes to C3' endo in dG residues. The structures of C2' endo, *anti* dG and C3' endo, *syn* dG are shown in Figure 20. The zigzag phosphate backbone is the result of the alternating C2' endo conformation in pyrimidines and C3' endo conformation in purines. One major consequence of the change in sugar pucker is the effect this has on the distance between the phosphate groups attached to the C5' and C3' ribose positions (Figure 20). Changing the sugar pucker from C2' endo to C3' endo reduces this distance dramatically. This shortening has the result of pulling the base away from the center of a B-DNA helix and closer to the phosphate backbone, which is on the outside of the DNA helix. The other significant change occurring in deoxyguanosine in Z-DNA is a 180° rotation around the glycosidic bond. The configuration goes from the *anti* configuration found in B-form (and A-form) DNA to the *syn* configuration in Z-DNA.

A major structural rearrangement required for the B-DNA to Z-DNA transition is a flipping or inversion of the bases relative to the helix axis. Within a region of Z-DNA the bases flip over 180°. This flipping is accompanied by the rotation of the glycosidic bond of the purines from *anti* to *syn* and the sugar pucker change.

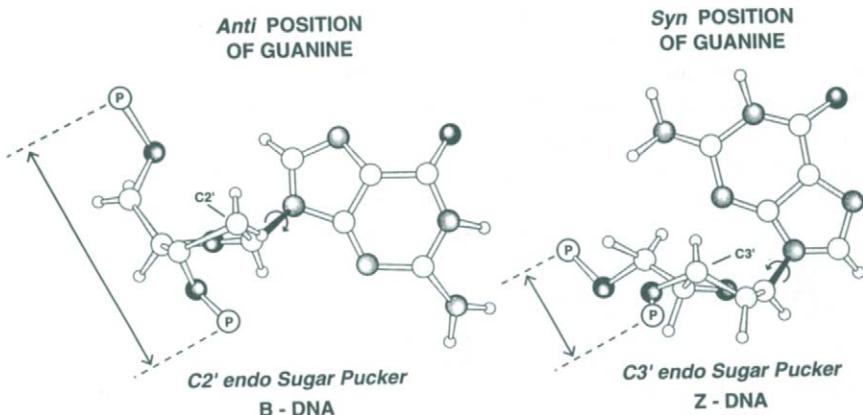


Figure 20. Conformations of guanosine in B-DNA and Z-DNA. B-DNA. The N9—C1' glycosidic bond (shown in black) is in the *anti* position, and the sugar is in the C2' endo pucker conformation. Z-DNA. The glycosidic bond is in the *syn* position (rotated 180° with respect to the *anti* position), and the deoxyribose is in the C3' endo pucker conformation. (Figure modified from Rich et al.³⁴⁸ with permission.) Reproduced with permission from the *Annual Review of Biochemistry*, Volume 53. ©1984 by Annual Reviews, Inc.

For the pyrimidine nucleotides in Z-DNA, the sugar accompanies the base in its 180° rotation.

2. Base Stacking and Positioning in the Z-DNA Helix

In B-form each DNA base pair is oriented in the helix similarly with respect to its adjacent base pairs, such that all dinucleotide base pairs are structurally similar. Another prominent feature of B-DNA is that the base pairs are stacked regularly on top of one another. As shown in Figure 21, there is reasonable overlap between the GpC dinucleotides. There is somewhat less physical overlap between the bases in the CpG stack. In B-form DNA the bases are hydrogen bonded in a position almost perfectly centered within the double helix (the dot in Figure 21 denotes the center of the DNA helix). As illustrated in Figure 22, the bases in B-form DNA stack into a cylinder within the center of the double helix, with the sugars and phosphates on the outside of the helix.

At the base pair level, another major difference between B-DNA and Z-DNA is that in Z-DNA the base pairs are displaced, or sheared, with respect to their stacking positions. The orderly stacking in the center of the helix found in B-DNA is nonexistent in Z-DNA. Analysis of the CpG dinucleotide in Figure 21 shows that the two cytosines are stacked, but the guanines are positioned under and over the ribose sugars. Thus, in the CpG dinucleotide the bases are sheared by being effectively pulled away from each other with respect to the center of the helix. In the GpC dinucleotide in Z-DNA, the bases are relatively well stacked.

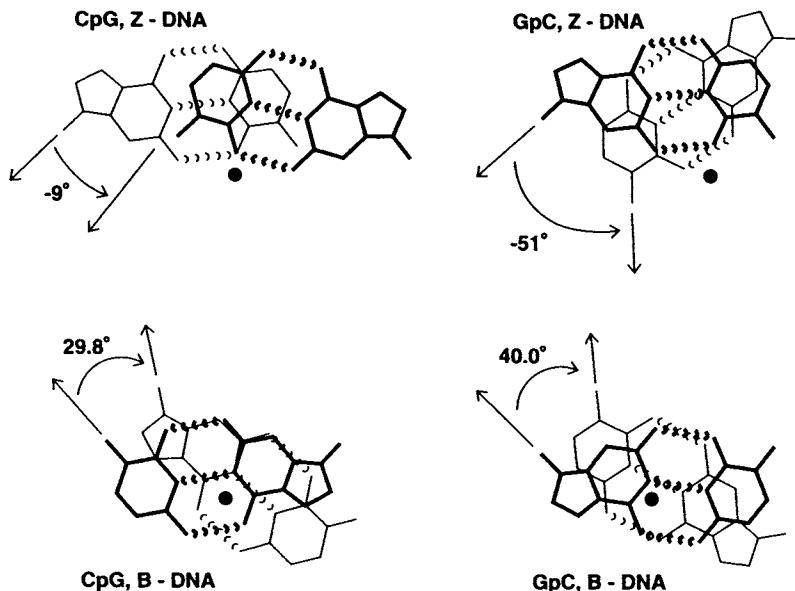


Figure 21. Base stacking in Z-DNA and B-DNA. The CpG and GpC base pairs are shown in Z-DNA and B-DNA. The bases in bold are stacked above the bases denoted by the thin lines. Bases are stacked well in B-DNA, especially in the GpC dinucleotide. Poor stacking overlap occurs in the CpG dinucleotide in Z-DNA. The twist angles shown are positive for B-DNA and negative for Z-DNA. (Figure modified from Rich et al.³⁴⁸ with permission.) Reproduced with permission from the *Annual Review of Biochemistry*, Volume 53. ©1984 by Annual Reviews, Inc.

The position of bases within the helix also distinguishes DNA from B-DNA. In B-DNA the bases form a cylinder within the double helix, with the hydrogen bonds at the center of the helix. However, as shown in Figures 21 and 24, the bases in Z-DNA are positioned toward the outside of the Z-DNA helix such that neither the bases nor the hydrogen bonds overlap the helix. In B-DNA, the bases are protected from the solvent by their location at the center of the helix. In Z-DNA, certain ring positions are much more chemically reactive than in B-DNA. For example, the N7 and C8 positions of guanine are exposed to the solvent in Z-DNA (Figure 22).

3. A Repeating Dinucleotide in Z-DNA

Because all base pairs are positioned similarly with respect to other base pairs, B-form DNA has a base pair repeating unit of 1. In Z-DNA the repeating base pair unit is 2. The CpG dinucleotide is very different from GpC. There are 12 bp/turn

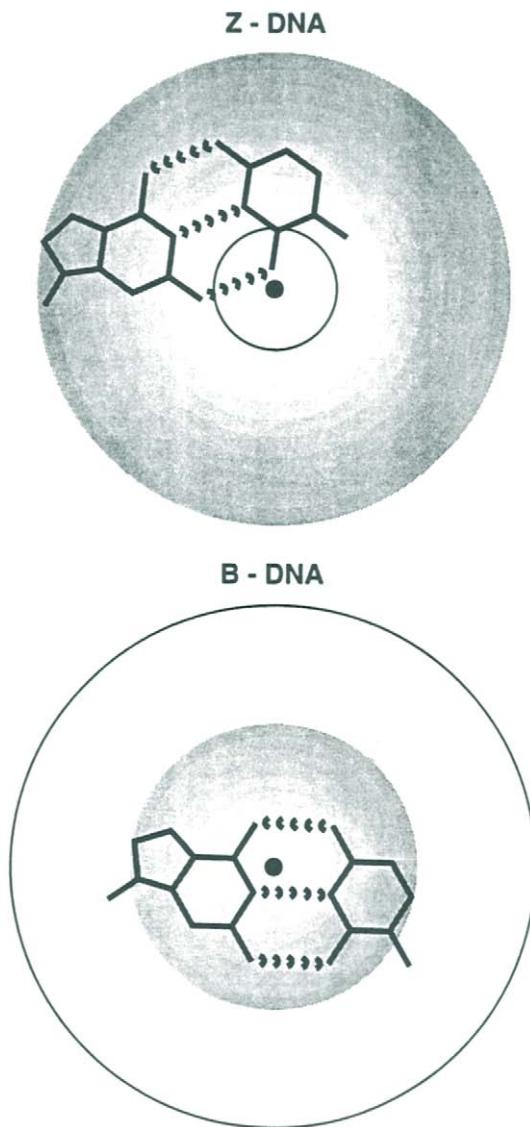


Figure 22. Location of base pairs within the B-DNA and Z-DNA helix. End-on views of the location of a G-C base pair in the Z-DNA and B-DNA helices. The shaded regions indicate the positions the base pairs occupy within the helical cylinder. The bases are centrally stacked in B-DNA, whereas they are organized toward the outside of the helix in Z-DNA.

of Z-DNA, which should produce an average helix twist of -30° (the value is negative, to denote the left-handed rotation in Z-DNA compared to a positive rotation found in right-handed helices). However, the twist angle for CpG is only -9° , whereas the twist angle for GpC is -51° (Figure 21). Consequently, in contrast to B-DNA, each base pair in Z-DNA is not oriented in the helix in a similar fashion with respect to the adjacent base pairs. However, every two base pairs are oriented within the helix in a fashion identical to the next two adjacent base pairs. The twist angle for a two base pairs relative to the next two base pairs is -60° , producing a helix repeating unit of 2.

4. The B-Z Junction

When a region of Z-DNA exists within a larger B-form DNA molecule, there must be a junction between the right- and left-handed helices. The best estimation of the physical structure of a B-Z junction is that it consists of a region of probably three or four unpaired or weakly paired bases. There is likely not a sharp transition from the left- to right-handed helix, but a region of a few base pairs that is partially unwound. The selective or differential reactivity of many chemicals at the B-Z junctions has provided sensitive assays for Z-DNA.

C. Formation and Stability of Z-DNA

1. The Sequence Requirement for Z-DNA

Z-DNA can form in regions of alternating purine-pyrimidine sequence, with $(GC)_n$ sequences forming Z-DNA most easily. $(GT)_n$ sequences also form Z-DNA, but they require a greater stabilization energy for formation than $(GC)_n$. $(AT)_n$ generally does not form Z-DNA, since $(AT)_n$ easily forms cruciforms. $(AT)_n$ can form Z-DNA under two special conditions. First, up to 10 alternating A-T base pairs must be embedded within a $(GC)_n$ or $(GT)_n$ region to form Z-DNA.³⁴⁹ Second, $(AT)_n$ can form left-handed DNA at high negative supercoiling or in high salt, and in the presence of $NiCl_2$.³⁵⁰

2. Chemical Modifications Stabilize Z-DNA

Certain chemical modifications of DNA will drive the B-DNA \rightleftharpoons Z-DNA equilibrium in favor of Z-DNA. Bromination at the C5 position of cytosine or the C8 position of guanine allows the stabilization of Z-DNA at lower salt concentrations than required to stabilize a nonbrominated polymer. The addition of a bulky group at the C8 position of guanine favors the *syn* conformation, which places this group on the outside of the Z-DNA helix (see Figure 22). In the *anti* conformation, the bulky group would be inside the cylindrical axis of the DNA near the phosphate backbone. Methylation of the N7 position of guanine also favors the formation of Z-DNA. Bromination and, especially, methylation (or ethylation) of DNA from

environmental mutagens will affect the stability of the B or Z conformation of the DNA double helix.

In bacterial and eukaryotic DNA many bases are methylated. Methylation is used in bacteria for restriction modification systems and the methyl-directed mismatch repair system. In eukaryotic cells, the C5 position of cytosine in the CpG dinucleotide is frequently methylated, as a general mechanism preventing transcriptional activity from large regions of eukaryotic chromosomes.

3. The Effect of Cations on the B-to-Z Equilibrium

The formation of Z-DNA was first observed in very high concentrations of salt (3 M NaCl). The distance between the negatively charged phosphates is much closer in Z-DNA than in B-DNA. Therefore, Z-DNA will be destabilized by the charge-charge repulsion of the negative charges on the phosphates. High concentrations of monovalent cations (Na^+ , K^+ , Rb^+ , Cs^+ , Li^+) can be required to shield the negative charges and stabilize the Z-conformation. Divalent cations (for example, Mg^{2+} , Mn^{2+} , and Co^{2+}) are much more effective at shielding negative charges than are monovalent cations and will stabilize Z-DNA at much lower concentrations. Co^{3+} will stabilize Z-DNA at millimolar concentrations. In addition, spermine, spermidine, and modified polyamines will stabilize Z-DNA.

4. DNA Supercoiling Stabilizes Z-DNA

In the early 1980s, experiments from the laboratories of Wells, Rich, and Wang demonstrated the role of supercoiling in Z-DNA formation.^{351,352} Klysik et al.³⁵³ showed that when a plasmid containing the sequence $(\text{dCdG})_n \cdot (\text{dCdG})_n$ was analyzed by agarose gel electrophoresis in the presence of 4 M salt, relaxation indicative of the formation of Z-DNA was observed. Subsequently, Stirdvant et al.³⁵⁴ and Peck et al.³⁵² showed that as the level of negative supercoiling increased, the concentration of salt required to drive the Z-DNA transition decreased. In fact, the level of negative superhelical energy necessary to drive the B-to-Z transition at physiological ionic strengths was well within the level of supercoiling found in DNA purified from cells. $(\text{dCdA})_n \cdot (\text{dGdT})_n$ also forms Z-DNA in supercoiled plasmids.^{355,356} However, $(\text{GT})_n$ Z-DNA-forming sequences require a higher level of negative supercoiling to form Z-DNA than a $(\text{GC})_n$ tract of equal size. The level of negative supercoiling required to form Z-DNA is a function of length of the alternating purine-pyrimidine tract. As a general rule, the longer the Z-DNA-forming sequence, the less negative supercoiling required to drive the B-to-Z transition. Table 4 shows the level of negative supercoiling required to form Z-DNA for a number of Z-DNA forming sequences.

DNA supercoiling is one physiological condition that can drive the formation of Z-DNA. Z-DNA is quite stable in supercoiled DNA, since the formation of Z-DNA effectively unwinds DNA, resulting in the relaxation of supercoils. As dis-

Table 4. Supercoiled Density Dependence for Z-DNA Formation

Z-DNA Sequence	σ_c	References
(CG) ₁₆	-0.031	Peck et al. ³⁵²
(CG) ₈ GG(CG) ₇	-0.034	Ellison et al. ³⁵⁷
GCGCGCGAGCGCGCGCGCTCGCGCG	-0.042	Ellison et al. ³⁵⁷
CG(TG) ₂₀ AATT(CA) ₂₀ CG	-0.032	Blaho et al. ³⁵⁸
(CG) ₁₃ AATT(CG) ₁₃	-0.025	Blaho et al. ³⁵⁸
(CG) ₆ TA(CG) ₆	-0.042	Sinden and Kochel ³⁵⁹
(TG) ₆ TA(TG) ₆	-0.057	Kochel and Sinden ³⁶⁰
(TG) ₂₆	-0.047	Haniford and Pulleyblank ³⁵⁵

cussed above, conditions that unwind DNA are thermodynamically favored. This situation is analogous to that described in the preceding section for cruciforms and the formation of intramolecular triple-stranded DNA discussed in the following section. When a 12-bp region of B-DNA undergoes a transition from a right-handed helix to one complete left-handed helical turn, the number of twists in the DNA will decrease by 2.14 (1.14 from the removal of B-DNA ($12 \div 10.5$); plus 1 from the formation of Z-DNA). In terms of 10 bp units, the formation of Z-DNA removes 1.78 turns for every 10 bp that form Z-DNA ($[10 \div 10.5] + [10 \div 12]$).

5. Protein-Induced Formation of Z-DNA

The widespread occurrence of alternating purine-pyrimidine sequences in eukaryotic DNA, especially alternating (GT) sequences, has led to speculation that Z-DNA- and Z-DNA-binding proteins must exist. The identification of proteins that bind Z-DNA might implicate a role for the left handed conformation of DNA in biology. Proteins that bind to Z-DNA have been identified from many organisms, including *E. coli*, *Drosophila*, and higher eukaryotes. It is not yet known with certainty, however, if any Z-DNA-binding protein has a biological role that actually involves binding to left-handed DNA.

The first example of a protein that could recognize and specifically bind to Z-DNA was a Z-DNA antibody. Anti-Z-DNA antibodies recognize left-handed Z-DNA but not right-handed DNA. Lafer et al.³⁶¹ showed that some antibodies could induce the formation of Z-DNA and stabilize Z-DNA by binding the left-handed helix in relaxed (linearized) plasmid DNA. This was a significant finding, since it provided an example of a protein that could affect an equilibrium between two alternative helical forms of DNA.

D. Assays for Z-DNA

The major structural differences between Z-DNA and B-DNA allow many physical and chemical approaches to be used as assays for these conformations. Many of these approaches are also applicable for cruciform and triplex DNA structures. Table 5 describes several assays for Z-DNA.

Table 5. Assays for Z-DNA

Assay	Description	Representative Reference
Z-DNA antibodies	Many polyclonal and monoclonal antibodies have been raised against Z-DNA. Binding can be detected by a variety of methods, including the binding of the antibody-DNA complex to nitrocellulose.	Lafer et al. ³⁶² Moller et al. ³⁶³
DNA topological assays	The formation of Z-DNA within a plasmid results in the relaxation of about 1.78 negative supercoils for every 10 bp of B-DNA that form Z-DNA. This relaxation can be detected on agarose (especially two-dimensional) gels.	Singleton et al. ³⁵¹ Haniford and Pulleyblank ³⁵⁵ Peck and Wang ³⁶⁴ Nordheim et al. ³⁶⁵
Chemical probes of Z-DNA	Many chemicals (such as diethylpyrocarbonate) react preferentially with the N7 and C8 positions of guanine when in the Z-DNA helix. Chemicals that specifically react with unpaired bases (bromoacetaldehyde (BAA), chloroacetaldehyde (CAA), osmium tetroxide (OsO_4), hydroxylamine, aminofluorene derivatives) react at the B-Z junctions. Sites of chemical modification can be mapped by S1 nuclease cleavage or primer extension analysis.	Johnston and Rich ³⁶⁶ McLean et al. ³⁶⁷ Rio and Leng ³⁶⁸ Kochel and Sinden ³⁶⁹ Hoepfner and Sinden ³⁷⁰
Restriction/modification assays	Psoralen and other intercalating drugs bind poorly to Z-DNA. Psoralen photobinds in a hypersensitive fashion to certain B-Z junctions. These characteristic differences provide an assay for Z-DNA. Psoralen photobinding can be measured with base pair resolution by an exonuclease III mapping procedure or by primer extension analysis.	Azorin et al. ³⁷¹ Singleton et al. ³⁷² Zacharias et al. ³⁷³ Vardimon and Rich ³⁷⁴

E. Z-DNA *In Vivo*

Z-DNA has been demonstrated in living cells by a number of different chemical, enzymatic, and physical approaches. One of the first *in vivo* assays for Z-DNA involved the analysis of psoralen cross-linking and photobinding to the Z-DNA forming sequence (CG)₆TA(CG)₆. The pattern of photobinding to this sequence in living cells was consistent with an *in vitro* superhelical density of $\sigma = -0.035$.^{359,375} Jaworski et al.³⁷⁶ provided evidence for Z-DNA by using an *Eco*RI methylase (MEcoRI) assay, in which the *Eco*RI site (GAATTC) was not methylated when it existed within a Z-DNA region. Haniford and Pulleyblank^{376a} used a linking number assay to provide evidence for the existence of Z-DNA *in vivo*. A bimodal distribution of supercoiled plasmid containing a (GC)_n Z-forming insert was not observed in cells under normal growth conditions, but was observed

in cells treated with chloramphenicol, which can lead to an increase in DNA supercoiling and thus drive Z-DNA formation. Several chemical probes have also been used to detect Z-DNA *in vivo*. Rahmouni and Wells³⁷⁷ applied osmium tetroxide (OsO_4), a chemical probe for unpaired pyrimidines, to detect Z-DNA *in situ*. Thymines within the EcoRI sites (GAATTC) between two $(\text{CG})_n$ blocks or flanking a $(\text{CG})_n$ block were hypersensitive to OsO_4 . In these experiments, reactivity was not observed with $(\text{CG})_5$ Z-DNA-forming sequences, but reactivity was detected with $(\text{CG})_6$ and longer tracts, indicative of the existence of Z-DNA.

Anti-Z-DNA antibodies have also been used extensively as probes for Z-DNA in eukaryotic cells. Nordheim et al.³⁷⁸ were the first to use Z-DNA antibodies to detect Z-DNA in eukaryotic cells in the polytene chromosomes of the *Drosophila* salivary gland. Because the harsh conditions initially required for the antibody binding may introduce negative supercoiling and thus Z-DNA formation,³⁷⁹ several approaches have been developed to maintain as natural a chromosomal state as possible for antibody binding studies. Jackson and Cook³⁸⁰ have shown that gentle purification of nuclei from cells can result in changes in the organization of DNA in chromosomes. Therefore, a protocol involving encasing eukaryotic cells in agarose beads was developed by Jackson and Cook. Cells are gently lysed in the beads, and the cytoplasm and membrane fragments can be washed out of the beads. The nucleus remains and is permeable to antibodies and certain restriction enzymes and nucleases. Using this approach, Whittig et al.^{193,194} have presented evidence consistent with the existence of Z-DNA in eukaryotic cells. The level of antibody binding suggested the existence of Z-DNA in every 100 kb of DNA. To demonstrate that Z-DNA was dependent on unrestrained supercoiling, nuclei in the agarose beads were treated with DNase I to introduce nicks into DNA. This reduced antibody binding. Moreover, when a topoisomerase I inhibitor was added, a higher level of anti-Z-DNA antibody binding was observed. These results are consistent with the interpretation that the formation of Z-DNA in eukaryotic cells is a DNA supercoiling-dependent process.

F. Possible Biological Functions of Z-DNA

Z-DNA may play regulatory roles in gene expression, DNA replication, or genetic recombination. Although a definitive role for Z-DNA remains to be established, this discussion will enumerate some of the possibilities. Many of the roles of Z-DNA are similar to those described below for triplex DNA.

The formation of Z-DNA could alter the level of supercoiling and possibly control the level of expression. In bacterial cells, gene expression can be regulated by the level of unrestrained supercoiling in DNA. Although many other factors are important, the expression of certain genes is regulated by supercoiling. The topoisomerase I and DNA gyrase genes of *E. coli* are two such examples of genes regulated by DNA supercoiling. Unrestrained supercoiling can exist at active gene regions, and the energy from supercoiling may facilitate RNA polymerase binding. The level of

unrestrained supercoiling will decrease within a topological domain as a region undergoes a transition from B-DNA to Z-DNA. If a region of Z-DNA returns to the B-form, the level of negative supercoiling will increase. Since a certain level of negative supercoiling may be required for optimal gene expression, the ability to alter supercoiling levels by B-to-Z or Z-to-B transitions provides one mechanism for gene regulation. DNA supercoiling also acts to compact chromosomes, which will increase the local concentration of the DNA. A change in the concentration of DNA can influence the binding of regulatory proteins. The very different helix structures of B-DNA and Z-DNA clearly provide different signals to DNA-binding proteins. The formation of B-DNA or Z-DNA within a particular DNA-binding site may present a defined substrate for specific proteins designed to recognize a particular DNA helix. Differential protein binding to the B- or Z-form of a helix may be required for gene expression acting as a switch in a positive or a negative fashion.

Several different Z-DNA-forming sequences have been cloned into the regulatory regions of several genes to examine the influence of a Z-DNA-forming sequence on transcription. To date there is no consensus for the role of Z-DNA in gene regulation. Z-DNA can have a positive, negative, or no effect on gene expression.^{381–384} This result may not be unexpected, given the complexity of eukaryotic gene regulation, including the multiple transcription factors that are likely to be associated with each promoter. Each promoter certainly binds many different transcription factors, and their binding may be differentially influenced by B- or Z-DNA. The existence of Z-DNA may also affect nucleosome positioning on promoters, which could strongly influence gene expression.

Z-DNA may be involved in genetic recombination, which may involve the transient formation of a left-handed helix (for a review see Blaho and Wells).³⁸⁵ When a single strand begins base pairing with a homologous duplex, a single turn of left-handed DNA may form for every helical turn of right-handed DNA formed. Within the base paired region, sequences with alternating purine-pyrimidine symmetry would adopt a left-handed helix most easily. DNA adjacent to a Z-DNA-forming region may preferentially be recombinagenic. In support of this idea, many DNA sequences with the potential to form Z-DNA have been associated with sites of genetic recombination.^{386–391} Additional evidence for Z-DNA in recombination comes from the observation that many proteins involved in recombination bind to Z-DNA. RecA (from *E. coli*) and RecI (from *Ustilago*) facilitate duplex alignment and strand exchange. These proteins bind preferentially to Z-DNA.^{385,390,392–394}

Z-DNA binding proteins have been purified from many different organisms, including *E. coli*, *Drosophila*, yeast, nematodes, chicken, frog, bull testis, and human cells,^{246,375,395–400} although, for the most part, the role of the proteins in binding Z-DNA is not understood. *Drosophila* topoisomerase II binds preferentially to Z-DNA.³³⁸ Moreover, binding of GTP to the enzyme weakens its affinity for B-DNA while increasing the binding to Z-DNA. In negatively supercoiled DNA regions, such as genes primed for or active in transcription, Z-DNA formation may act to facilitate topoisomerase II binding.

VI. TRIPLE-STRANDED NUCLEIC ACIDS

A. Introduction

The triple-stranded helices (triplexes) of nucleic acids were first described in 1957⁴⁰¹ and have been studied for more than 30 years.^{1,2,402–406} Intermolecular triplexes attract much attention because of their potential therapeutic and biotechnological applications.^{2,407–410} Intramolecular triplex/single strand structures (H- and H*-DNA) are presumed to play important roles in DNA function (see refs. 1, 2, 404, and 406 for reviews).

B. The Structure of Triplex DNA

1. Sequence Requirements

A triple helix consists of a duplex, in which the bases are paired via Watson-Crick hydrogen bonds, and a third strand, the bases of which form hydrogen bonds with one base of each base pair of the duplex. After the formation of hydrogen bonds in Watson-Crick base pairs, purine bases have potential hydrogen-bonding donors and acceptors that can form two hydrogen bonds with incoming third bases. By contrast, each pyrimidine base already involved in the duplex can form only one additional hydrogen bond with incoming third bases. The hydrogen bonds of such a type are traditionally called Hoogsteen bonds, after their discovery in adenine-thymine cocrystals by Hoogsteen in 1959. To maximize the number of stabilizing hydrogen bonds, the third strand bases bind to the purine bases of the duplex. If these purines were randomly distributed in both duplex strands, this would result in an energetically unfavorable conformation of the sugar-phosphate backbone of the third strand due to base switching from the one duplex strand to the other, and in the lack of stabilizing stacking interactions in the third strand. Therefore, a duplex appropriate for stable triplex formation contains purine bases only in one strand. Thus the precondition of triplex formation is the presence of a homopyrimidine (Py) sequence in one strand of the duplex and a complementary homopurine (Pu) sequence in the opposite strand (Py-Pu tract).

Two hydrogen bonds with the duplex purine strand can be formed by both pyrimidine and purine bases of the third strand.^{2,411–413} Energetically stable triple helices are formed when the third strand is composed of either only pyrimidine bases (Py-Pu-Py type triplex) or mainly purine bases with a low proportion of pyrimidine bases (Py-Pu-Pu type triplex) (the third strands are shown in italics). Triplices can be also divided into intramolecular and intermolecular types (Figure 23).

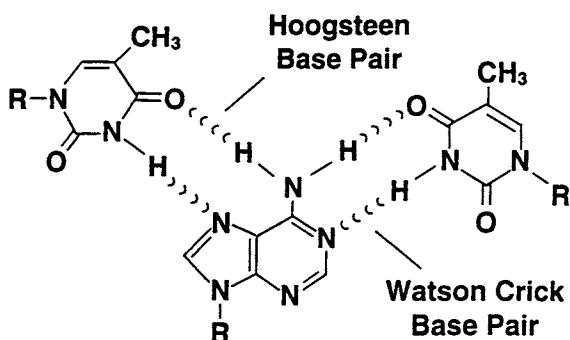
Typical intermolecular triplices can be formed when the polymeric or oligomeric third strand of appropriate sequence binds (through Hoogsteen bonding) to the double-stranded Py-Pu tract.^{401,414–416} A DNA sequence containing a Pu-Py tract of mirror symmetry can form an intramolecular H- or H*-DNA structure in which one half of a Pu-Py tract unwinds and then a Py (or Pu) strand bends around

A**Py•Pu•Pu****B****Py•Pu•Py**



Figure 23. Intermolecular and intramolecular triple helices. (A) An intermolecular Py-Pu-Pu triple helix is shown, with the polypurine third strand organized antiparallel with respect to the purine strand of the Watson-Crick duplex. (B) An intermolecular Py-Pu-Py triple helix is shown, with the polypyrimidine third strand organized parallel with respect to the purine strand of the Watson-Crick duplex. (C) An intramolecular triplex DNA is shown. As with intermolecular triplex DNA, the third strand lays in the major groove, whereas its complementary strand exists as a single strand. (Figure modified from Wells et al.⁵⁹⁴ with permission.)

A



Hoogsteen
Base pair

Watson Crick
Base pair

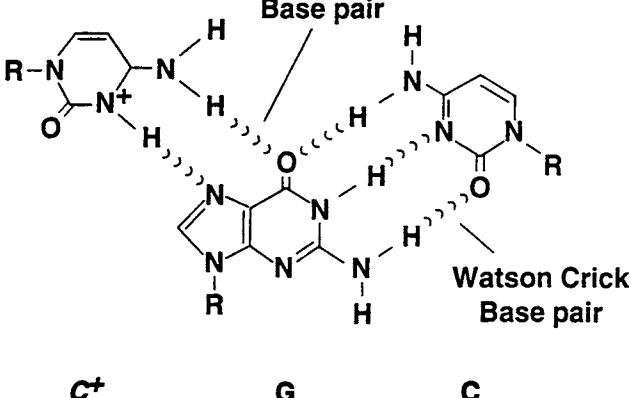
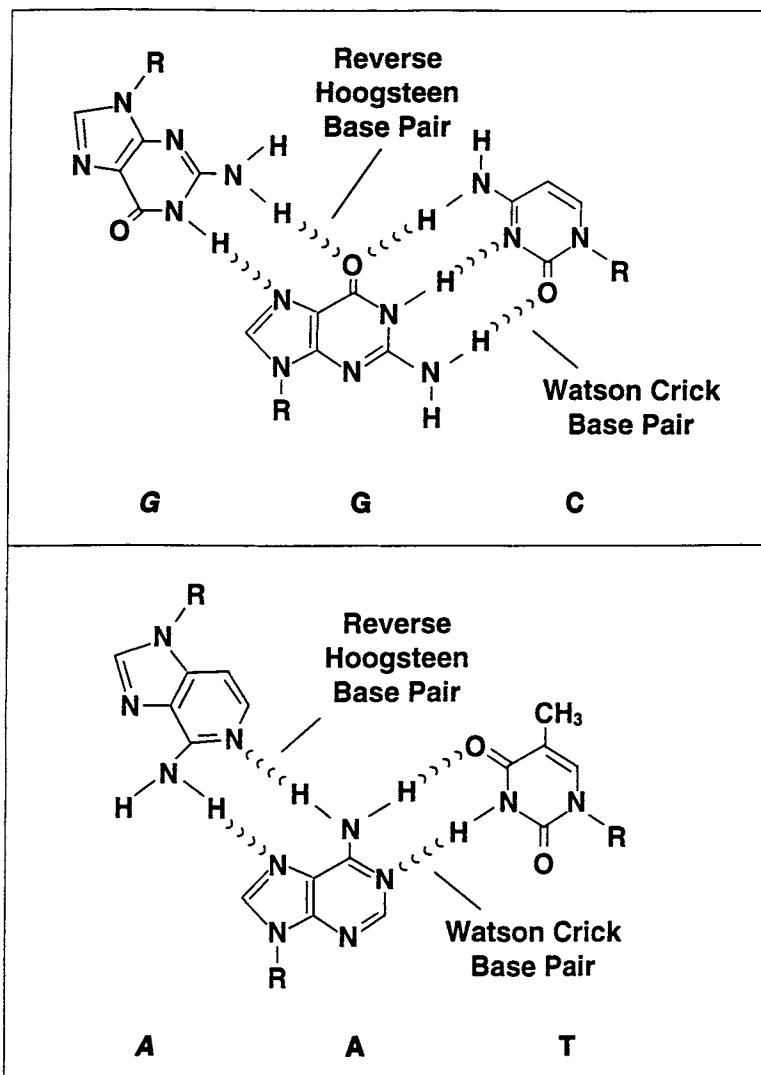


Figure 24. Base triads in triplex DNA. (A) The top panel shows the T-A-T triad, bottom panel the C-G-C⁺ triad. In the T-A-T triad, A forms both Watson-Crick and Hoogsteen base pairing to thymines. In the C⁺-G-C triad, G forms a Watson-Crick base pair and a Hoogsteen base pair with a protonated cytosine. In Py-Pu-Py triads the third base forms hydrogen bonds in a Hoogsteen configuration. (B) In the C-G-G triad the central G of the Watson-Crick base pair forms a Hoogsteen base pair with the third-strand G. In the T-A-A triad, the central A forms hydrogen bonds in a Hoogsteen base pair with an additional A. In Py-Pu-Pu triads, the third base forms hydrogen bonds in a reverse Hoogsteen configuration.

B



a point of symmetry and makes Hoogsteen hydrogen bonds with the Pu strand in another half of the Pu-Py tract. A former Pu (or Py) counterpart of the folded-back strand remains unpaired.^{417,418}

2. Isomorphism of Base Triads

Figure 24 shows stable base triads that can be formed with natural bases. Two configurations of third-strand bases relative to the Watson-Crick base pair are pos-

sible: Hoogsteen base pairing,⁶ which is realized in the Py·Pu·Py triplex, and reverse-Hoogsteen base pairing for the Py·Pu·Pu triplex, in which third-strand bases are rotated 180° relative to their positions in the previous scheme. Note that not all of the possible 16 Py·Pu·Py and 16 Pu·Pu·Pu triads⁴¹⁹ are stable, most of them would have one or zero Hoogsteen hydrogen bonds. Base triads in a Py·Pu·Py triplex (C·G·C⁺ and T·A·T) are isomorphous, that is, they can be superimposed one over another, so that the positions and orientations of corresponding glycosidic bonds, as well as positions of C1' atoms, practically coincide. The conformation of the sugar-phosphate backbone is regularly repeated along the third strand. The Py·Pu·Pu triplexes include C·G·G and T·A·A triads, which are not truly isomorphous; so the conformation of the backbone is not perfectly regular along the third strand. The Py·Pu·Pu triplex can also accommodate a number of T·A·T triads at the expense of some distortion in optimum base pairing and/or backbone conformation.^{405,420}

3. Major Groove Location and Antiparallel Orientation of the Third Strand

Base pairs of double-helical nucleic acid molecules retain hydrogen-bonding capabilities in both the minor and major grooves of the double helix (see, for review, Helene and Lancelot).⁴²¹ The third strand is placed in the major groove of the DNA, where purine bases may form two normal Hoogsteen hydrogen bonds (see, for a review, Cheng and Pettitt).⁴¹³ The major groove of DNA is deep enough and, upon third-strand binding, the resulting triplex has a diameter that is only a few angstroms larger than the 20-Å duplex diameter.⁴²²

An antiparallel orientation of similar strands and *anti* glycosidic bonds was observed in both the Py·Pu·Py and Py·Pu·Pu triplexes.^{2,405,423} (Some researchers call Py·Pu·Py triplexes parallel because the purine and the third strand are parallel, and Py·Pu·Pu triplexes antiparallel because the duplex purine strand and the third strand are antiparallel.)

4. Strand Geometry Is Possibly Intermediate between A- and B-Forms

Depending on the sequences involved and the type of third strand, DNA triplexes may have differences in overall structures, with the geometric parameters of constituent strands in triple helix close to the B-form or intermediate between the B- and A-forms.^{405,413,422,424–426}

C. The Formation and Stability of Triplex DNA

1. Kinetics of Triplex Formation

The formation of H-(H*)-DNA requires a denaturation bubble in the center of the mirror-repeated Py·Pu tract that allows the duplexes on either side to rotate

and fold back, providing the formation of the first triad.⁴²⁷ Subsequently, one-half of the Py-Pu tract hydrogen bonds to a single strand as it is released by the progressive unwinding of the other half of the Py-Pu tract. The rate of intramolecular triplex formation depends on the sequence of the Py-Pu tract, so that H-(H⁺)-DNA can form between 2 min and several hours, and can be increased by increasing the temperature.^{428,429} The reverse transition, from intramolecular triplex to duplex, takes less than several minutes.⁴²⁸ The association rate constants for intermolecular Py-Pu-Py triplexes were on the order of $10^3 \text{ M}^{-1}\text{s}^{-1}$, which is 10^3 times lower than the corresponding constant for duplex formation. The triplex dissociation rate constants were in the 10^{-5} to 10^{-4} s^{-1} range, resulting in a triplex lifetime of 1–10 hours at 37°C.^{430,431} Dissociation constants for the complex of 21–22-mer pyrimidine oligonucleotides with their DNA target may be as low as 1–10 nM.^{430–432}

2. Electrostatic Stabilization

The formation of triple-stranded nucleic acids requires a reduction of repulsion between negatively charged phosphate groups of the three strands. The repulsion between these phosphates might be screened by submolar concentrations of monovalent cations (e.g., Li⁺, Na⁺, K⁺).^{401,412,433} Much lower (millimolar) concentrations of divalent metal cations (e.g., Mg²⁺ or Mn²⁺), which bind more tightly to phosphates, stabilize triplex *in vitro*.⁴⁰¹ Submillimolar concentrations of polyamines, the distributed charges of which allow them to interact simultaneously with different nucleic acid strands, are effective in stabilizing triple-helical nucleic acids *in vitro*.^{434–438} Lysine-rich peptides as models of basic protein fragments have a triplex-stabilizing effect comparable to that of polyamines.⁴³⁹ If triplexes contain clusters of charged bases in third strands (e.g., protonated cytosines), the repulsion between these bases may partly destabilize triple helix.^{440,441} See Table 6 for a list of triplex-stabilizing factors.

3. Hoogsteen Hydrogen Bonds and Stacking Interactions

The sequence specificity of triplex formation is the major evidence of the importance of Hoogsteen hydrogen bonds. All stable triads have two Hoogsteen hydrogen bonds. The importance of two Hoogsteen hydrogen bonds explains the requirement of low pH for Py-Pu-Py triplexes where C-G-C⁺ triads form under acidic pH conditions because of protonation of the N3 of the third-strand cytosines (see, for a review, Frank-Kamenetskii).⁴⁴⁴ In many cases the Py-Pu-Pu triplexes are stabilized by divalent cations.^{418,445,456} Stronger Hoogsteen hydrogen bonds are possibly formed when divalent metal cations coordinate to purine bases in the third strand.²

Table 6. Triplex Stabilizing Factors

Type	Characterized <i>in vitro</i>	Availability <i>in vivo</i>	Reference
Electrostatic stabilization	Monovalent cations (Na^+ , K^+), up to 1 M	0.15 M	Krakauer and Sturtevant ⁴³³
	Divalent cations Mg^{2+} , 1–10 mM Zn^{2+} , 0.1–10 mM	Up to 1 mM Bound in metalloenzymes	Felsenfeld and Miles ⁴¹¹ Darnell et al. ⁴⁴²
	Polyamines: putrescine (2+) >1 mM Spermidine (3+) > 0.1 mM	May be as high as 1 mM, mostly in a macromolecule-bound form. Free polyamines in a micromolar range	Hampel et al. ⁴³⁵ Singleton and Dervan ⁴³⁷ Thomas and Thomas ⁴³⁸ Sarhan and Seiler ^{442a}
	Spermine (4+) > 0.01 mM		Davis et al. ⁴⁴³
	Basic polypeptides, 0.01–1 mM	Probably as protein fragments	Potaman and Sinden ⁴³⁹
Hoogsteen hydrogen bonds	Similar or slightly weaker than Watson–Crick Low pH for C·G·C ⁺ and C·G·A ⁺ triads Transition metal cations for Py·Pu·Pu triplex	No No	Cheng and Pettitt ⁴¹³ Frank-Kamenetskii ⁴⁴⁴ Bernues et al. ⁴⁴⁵
Stacking	Stabilizing contribution similar to duplex DNA	Yes	Roberts and Crothers ⁴⁴⁶ Cheng and Pettitt ⁴¹³
Hydration	Spines of hydration in all three grooves	Yes	Radhakrishnan and Patel ⁴⁰⁵ Weerasinghe et al. ⁴⁴⁷

Supercoiling for H (H*) DNA	Negative supercoiling dependence on Py-Pu length and pH	Static, transcriptionally and chromatin rearrangement-induced supercoiling	Lyamichev et al. ⁴⁴⁸ Kohwi and Kohwi-Shigematsu ⁴¹⁸ Sinden ¹
Length of third strand	≥ 6 nt for H-DNA, ≥ 10 nt for intermolecular	One H-DNA motif in 50,000 bp of human DNA One triplex-forming sequence in 300–500 bp	Lyamichev et al. ⁴⁴⁹ Schroth and Ho ⁴⁵⁰ Horne and Dervan ⁴⁵¹ Behe ^{451a}
Ligands	Intercalating agents		Pilch and Breslauer ⁴⁵² Duval-Valentin et al. ⁴⁵³ Wilson et al. ⁴⁵⁴ Thuong and Helene ⁴⁰³
Covalently bound agents	Intercalators, chemically and photochemically cross-linking agents		Thuong and Helene ⁴⁰³ Soyfer and Potaman ²
Modified third strand	Non-ionic backbones, modified nucleobases		Thuong and Helene ⁴⁰³ Soyfer and Potaman ² Nielsen et al. ⁴⁵⁵

Experimental studies of mismatches and bulges in triplex structures show the general importance of stacking interactions.^{431,446,457,458} The free energy penalty for introducing a single mismatch ranges from 2.5 to 6 kcal/mol, which is close to the corresponding values for DNA and RNA.

4. *Hydration Effects*

The NMR and molecular dynamics simulation data show that long-lived water molecules with lifetimes of >1 nsec are immobilized in all three grooves of the Py·Pu·Py and Py·Pu·Pu triple helices.^{405,447} Water molecules may stabilize the triplex structure by solvating the complex, by screening repulsive electrostatic interactions between phosphate groups across the narrow groove between the third strand and duplex purine strand of the Py·Pu·Py triplex, or by bridging polar groups belonging to different strands.⁴⁰⁵

5. *Requirement of Supercoiling for Intramolecular Triplex*

Relaxation of some torsional stress in closed circular DNA may result from some non-B-form DNA structure (reviewed in refs. 1, 177, 459). Therefore, increased levels of DNA supercoiling can provide the formation and stabilization of intramolecular Py·Pu·Py triplex (H-DNA)^{427,448,460} or Py·Pu·Pu triplex (H*-DNA).⁴¹⁸ Increased DNA superhelicity and low buffer pH may substitute for each other, and the superhelical tension necessary for H-DNA formation depends linearly on increasing pH.⁴⁴⁸

6. *Length Dependence*

Increasing the length of the Py·Pu tracts facilitates the formation of triplexes—the longer the lengths of Pu-Py tracts, the less the negative superhelix densities are required to induce the intramolecular triplex.^{427,449,460} A theoretical lower limit of 15 bp for a Py·Pu tract was estimated for H-DNA formation.⁴⁴⁹ At Py·Pu tracts that are several dozen nucleotides long, multiple conformers may form.^{429,461} The length dependence for intermolecular Py·Pu·Py triplex showed an increase in triplex stability with increasing oligomer length.^{462,463} In the case of Py·Pu·Pu triplex, the length of the Pu third strand seems to have an optimum because of an imperfect fit to the major groove of the duplex.^{412,464,465}

D. Assays for Triplex DNA

A number of physical and chemical techniques can be used as assays for H- and H*-DNA *in vitro* and *in vivo* (Table 7).

Table 7. Assays for H- and H*-DNA

Assay	Description	References
Two-dimensional agarose gel	The formation of H-DNA results in the relaxation of one negative supercoil for every 10.5 bp of B-DNA that forms H-DNA.	Lyamichev et al. ⁴⁴⁸
Chemical probes	Chemicals that specifically react with unpaired bases react at the tip of triple helix and unpaired half of purine or pyrimidine strand. Dimethyl sulfate reacts with the N7 of guanines, and protection of guanines from modification indicates the bases involved in the triple helix. Similarly, pyrimidines protected from photochemical modification map to a triplex-forming sequence. The pattern of 4,5',8-trimethylpsoralen photobinding to AT and TA dinucleotides in duplex provides a quantitative measure of triplex formation.	Kohwi and Kohwi-Shigematsu ⁴¹⁸ Johnston ⁴⁶⁶ Hanvey et al. ⁴⁶⁷ Lyamichev et al. ⁴⁶⁸ Ussery and Sinden ⁴⁶⁹
Enzymatic assays	Single-strand-specific nucleases recognize and cut unpaired regions in H- and H*-DNA. Inability of restriction enzyme to cut DNA indicates the presence of triple-stranded region.	Lyamichev et al. ⁴⁴⁸ Hanvey et al. ⁴⁶⁷

E. Triplex DNA *In Vivo*

1. Search for Triples in Cells

Several attempts have been made to determine whether the Py-Pu sequences do really form triple-stranded structures in the cell. Information from these studies may be separated into results from direct determinations and indirect indications of the triplices *in vivo*.

Immunological assays. To directly probe the triple-stranded structures in chromosomes, monoclonal antibodies were produced by immunizing mice with poly[d(Tm⁵C)]·poly[d(GA)]·poly[d(m⁵CT)] triplex, which is stable at neutral pH.^{470,471} These antibodies did not bind to calf thymus DNA or other non-Py-Pu DNAs, such as poly[d(TG)]·poly[d(CA)], and did not recognize Py-Pu DNAs containing m⁶A (e.g., poly[d(TC)]·poly[d(Gm⁶A)]) which cannot form a triplex since the methyl group at position 6 of adenine prevents Hoogsteen base-pairing. One type of monoclonal antibody was demonstrated by numerous criteria to be specific for the T·A·T-rich Py-Pu-Py triplex DNA, whereas another type was specific for poly[d(TC)]·poly[d(GA)]·poly[d(CT)] triplex. Chromosomes fixed in methanol/

acetic acid were stained by antibodies in the presence of *E. coli* DNA, but not in the presence of polymer triplex. To avoid an ambiguity in interpretation that could arise as the acid fixation itself can change the structure under study,³⁷⁹ nuclei were also stained after fixation in cold acetone.⁴⁷⁰ Unfixed, isolated mouse chromosomes also reacted positively with the antibody, particularly when they were gently decondensed by an exposure to low ionic conditions at neutral pH, indicating that fixation is not mandatory for antibody staining.¹⁹² Additional evidence that triplexes really do exist *in vivo* is provided by immunoblotting of triplexes in crude cell extracts.⁴⁷² Thus there is a growing body of immunological evidence that the triplexes are present in eukaryotic chromosomes.

Chemical probing. Direct chemical probing techniques were also used to detect triplexes in plasmid DNA in *E. coli*. The osmium/bipyridine modification pattern characteristic for H-DNA was observed in plasmid DNA after its host *E. coli* cells were preincubated in the pH 4.5 or 5.0 media.⁴⁷³ Upon cell incubation in the presence of Mg²⁺ and chloramphenicol at neutral pH, chloroacetaldehyde reactivity of DNA triplex-forming sequence (dG)₃₀ was similar to that for the C-G-G triplex *in vitro*.⁴⁷⁴ Glaser et al.,⁴⁷⁵ using diethylpyrocarbonate, failed to detect any modification pattern consistent with the presence of H-DNA, whereas in experiments *in vitro* they were able to detect it. High levels of DNA supercoiling and proper environmental conditions are the major limiting factors in the formation and detection of H- or H*-DNA structures.

Photochemical probing. Trimethylpsoralen photobinding to the Py-Pu tracts in plasmid DNA showed that the formation of H-DNA in *E. coli* cells was dependent on DNA superhelicity and extracellular pH.⁴⁶⁹ When cells were grown in K media, which acidifies over time down to pH 5, the trimethylpsoralen photobinding pattern was consistent with the presence of H-DNA. The use of topoisomerase I-mutant cells with a higher level of supercoiling *in vivo* was the triplex-promoting condition in these experiments.

Indirect assays. Several indirect studies addressed the possibility of triplex existence *in vivo*. The GATC site at the center of or adjacent to a Py-Pu mirror repeat was undermethylated in plasmid grown in *E. coli* strain JM101.⁴⁷⁶ This result could be explained by the participation of this Py-Pu tract in H-DNA *in vivo*, since *Dam* methylase methylates the GATC site when it is in the double-stranded B-DNA form, but not in an alternative non-B conformation. However, other unknown factors might be involved in *Dam* undermethylation as partial methylation of H-DNA *in vitro* was detected, and the GATC undermethylation was observed when the plasmid was grown in JM cells, but not when grown in other types of *E. coli* cells.

Deletion analysis of the Py-Pu tracts inserted in the tetracycline resistance gene⁴⁷⁷ showed a significant instability of those (longer) inserts that can form

intramolecular triplexes (H- or H*-DNA) *in vitro*. These data may reflect the existence and mutational role of triplexes *in vivo*, provided there is a satisfactory model for the underlying process.

Sarkar and Brahmachari⁴⁷⁸ showed that the Py-Pu tract cloned into the transcribed region of a bacterial gene significantly decreased the gene expression. Some unusual structure in the Py-Pu tract might be responsible for the premature transcriptional termination. Rao reported that the cloned d(GA)_n-d(TC)_n sequences that can potentially adopt triplex structures could slow down the DNA replication fork movement.⁴⁷⁹ In both of these cases the H-DNA was suggested to be responsible for preventing polymerases from progressing along the DNA template in the cellular system.

In summary, none of the triplex searches *in vivo* has been conclusive. This is not surprising, since if triplex DNA participates in gene functions it cannot exist all the time, at least under physiological conditions, because a truly stable structure cannot regulate a changing physiological state.⁴⁸⁰ It is hoped that further work will provide more convincing evidence of the existence of triplexes *in vivo*.

2. Factors that Could Be Responsible for Triplex Formation In Vivo

Supercoiling. Topoisomerases can create definite levels of (unrestrained) DNA supercoiling (see refs 1, 106, and 263 for reviews) in topological domains created by proteins and/or RNA molecules. Nucleosomal dissociation during transcription may convert a restrained supercoiling into an unrestrained supercoiling.¹ During the progression of RNA polymerase, a wave of negative supercoiling arises in DNA behind the RNA polymerase and positive supercoiling in front of the RNA polymerase (see Liu and Wang¹¹³ and Sinden¹ for more detail). A number of studies experimentally confirmed the difference in the level of supercoiling upstream and downstream of the transcribed site.^{134,260,262,377,481,482} An assortment of other mechanisms may create localized or transient torsional stress in eukaryotic DNA: binding of transcription factors or other proteins, activity of helix-tracking proteins, looping of DNA by protein binding at two distant locations, histone acetylation, and gyrase activity of topoisomerases.⁴⁸³

Cations. Upon triplex formation, the very high negative charge density that originates from the phosphate groups in three strands may be reduced by counterions (Table 6). Concentrations of metal cations in cells are far from stabilizing,⁴⁴² and metal cations may only be partially responsible for triplex formation and maintenance *in vivo*. Concentrations of polyamines in the nucleus may be as high as 5 mM,^{442a} where they are largely bound to nucleic acids and phospholipids.⁴⁴³ Thus polyamines might be a class of compounds which stabilize triplex DNA *in vivo*. Basic polypeptides stabilize triplex DNA *in vitro*.⁴³⁹ A similar stabilizing effect in cells might come from the basic protein fragments.

Proteins. The H-DNA structure formed through the interaction of two distant Py-Pu tracts of the same DNA duplex was proposed to be stabilized by single strand-binding proteins that could fix an unpaired single strand.⁴⁸⁵ About 20 proteins binding to Py-Pu sequences have been reported. Several examples include transcription factors such as Sp1⁴⁸⁵ and PuF,⁴⁸⁶ which bind to double-stranded DNA, and proteins that preferentially bind to homopyrimidine^{487–490} or homopurine single strands.^{491,492} In the absence of conclusive evidence on the role of Py-Pu tract-binding proteins, it is suggested that they may shift the equilibrium between the double-stranded conformation and H-DNA that has an unpaired strand.^{487,491,492} For example, such proteins may bind to and “capture” a single-stranded region resulting from Py-Pu tract breathing or denaturation under torsional stress.^{487,491}

Another option in H-DNA stabilization consists of a preferential protein interaction with the triple helix.⁴⁹³ The triple helix geometry or high negative charge density might be the features recognized by triplex-binding proteins. In support of this option, experiments with basic oligopeptides as reasonable models of surface-localized protein domains showed their triplex-stabilizing effect.⁴³⁹ However, there have been no direct data on triplex stabilization by proteins, and in some cases the experimental results were consistent with protein binding to Py-Pu tracts that had B-conformation but not H-DNA structure.^{475,494,495}

To summarize, there is an assortment of various factors that may induce and/or stabilize the triple-helical structures in the cell. When considered separately, these factors contribute to triplex stability *in vitro*. However, many of them are available in cells simultaneously, and their combined effects are not evident. For example, the stabilizing effects of cations of various valences interfere with each other, since these cations compete for the same binding sites.^{430,437} Clearly, more data are necessary to understand how triplexes could be formed and maintained in living cells.

F. Possible Biological Roles of H-(H*)-DNA

Py-Pu tracts capable of forming triple-stranded DNA structures have been found in the genomes of various organisms. In eukaryotes, the Py-Pu tracts a few dozen base pairs long constitute up to 1% of the entire genome.^{451a,496,497} Analysis of the human genome showed one H-DNA-forming sequence in every 50,000 bp.⁴⁵⁰ H-DNA has been suggested to play a role in key biological processes such as transcription, replication, recombination, and DNA condensation (see, for a review, Mirkin and Frank-Kamenetskii,⁴⁰⁴ Sinden,¹ Frank-Kamenetskii and Mirkin,⁴⁰⁶ and Soyfer and Potaman²).

1. Possible Regulation of Transcription

The Py-Pu tracts often occur in 5' flanking regions of eukaryotic genes, and a number of these tracts have been shown to be sensitive to the single-strand-spe-

cific nucleases (see Soyfer and Potaman² for a list of sequences). When cloned in plasmids, many of these Py-Pu tracts can adopt an H-DNA structure *in vitro*.^{387,448,475,495,498–501} H-DNA might influence the regulation of gene expression in a number of different ways: by affecting the level of DNA supercoiling in the topological domain in which it forms; by changing the local as well as the global structure of DNA; and by influencing nucleosome organization and nucleosome phasing from the triplex region.

Indirect influence of H-DNA on transcription. Generally, the efficiency of transcription increases with increasing superhelical density of the templates.¹⁰⁷ However, some bacterial and eukaryotic transcription systems require an optimum superhelical density.^{502–506} Thus H-DNA formation, partially relieving excessive superhelical tension, could serve to maintain the optimum template topology (as suggested in ref. 507).

Direct involvement of H-DNA in transcription. The formation of H-DNA can make the structure of the Py-Pu tract inappropriate for the transcription factor and subsequent RNA polymerase binding, thereby inhibiting gene expression (Figure 25A). For example, long triplex-forming (G)_n-(C)_n tracts placed 5' to the promoter inhibited transcription.⁵⁰⁸ The gel comigration data show that the (G)_n-(C)_n tracts can bind an activator protein present in human cells⁵⁰⁸ and chicken BPG1 protein.⁵⁰⁹ The local unwinding of DNA in the promoter sequence of the human Na,K-ATPase α 2 gene created by H-DNA formation extends into an adjacent TATA box, and thereby may down-regulate transcription by disrupting interactions between DNA and TATA-box-binding protein.⁵¹⁰

Structure-function analyses for many eukaryotic genes have demonstrated a stimulatory influence of the Py-Pu sequences on promoter function. Transcription efficiency was partially lost when the Py-Pu tracts were deleted from promoter regions of human epidermal growth factor receptor,⁵¹¹ c-myc,^{512,513} ets-2,⁵¹⁴ and decorin genes⁵¹⁵; rat neuronal cell adhesion molecule gene⁴⁹⁴; mouse c-Ki-ras⁵¹⁶ and transforming growth factor β 3⁵¹⁷ genes; and *Drosophila* hsp26⁴⁷⁵ and actin⁵¹⁸ genes. Cloning of the H-forming Py-Pu tract upstream of the β -lactamase promoter stimulated transcription compared with the plasmid lacking the Py-Pu tract.⁵⁰⁷ Triplex-distorting variations in the sequence reduced the transcriptional efficiency. Similarly, mutations in the repeating c-myc sequence motif (ACCCTC-CCC)₄, which would result in more mismatches within the suggested triplex, led to reduced transcriptional activity of the gene.⁵¹⁹ Two models of positive H-DNA influence on transcription may be consistent with these data.

A B-DNA-binding protein may interact with the promoter sequence as a repressor, preventing RNA polymerase binding (Figure 25B). When H-DNA forms, the repressor protein can no longer bind to the promoter sequence, allowing RNA polymerase access to the promoter sequence. The formation of H-DNA at the repressor binding site might be similar to the action of inducer protein, which

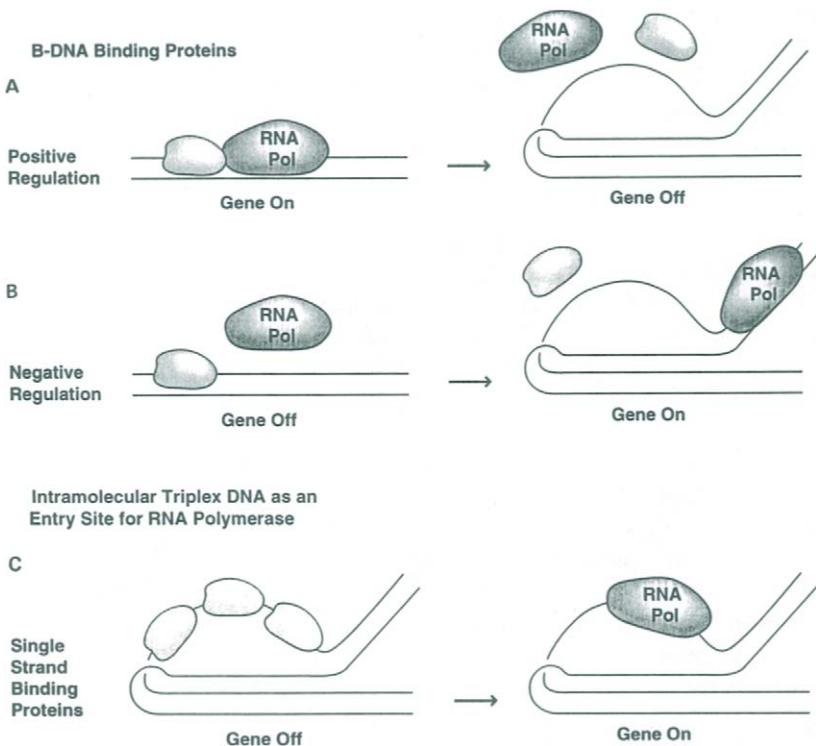


Figure 25. Models for intramolecular triplex involvement in gene regulation. (A) A positive acting B-DNA-binding protein may promote binding of RNA polymerase and turn on transcription (left). Intramolecular triplex formation might prevent binding of the positive activator, thus preventing binding of RNA polymerase (right). (B) In the negative regulatory scheme, a B-DNA-binding repressor that prevents RNA polymerase binding to the promoter (left) may no longer bind to its recognition sequence when in the intramolecular triplex form (right). RNA polymerase may be able to bind to the unwound region and begin transcription. (C) Single-strand-binding proteins bound to the single-stranded loop of the intramolecular triplex structure might prevent RNA polymerase binding and thus transcription (left). The unwound single-strand region or the duplex-triplex junction may provide an entry site for RNA polymerase to the promoter (right) (as in model B).

serves to displace the repressor and facilitate RNA polymerase binding at the adjacent site. No experimental data are available to illustrate this attractive hypothetical option in the utilization of the H-DNA by living cells.

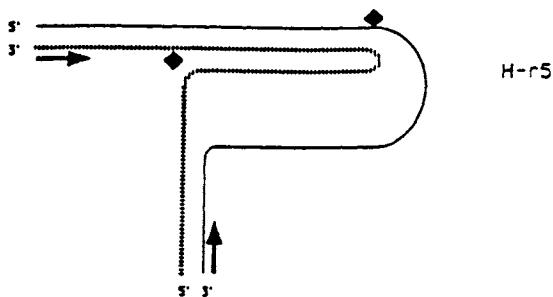
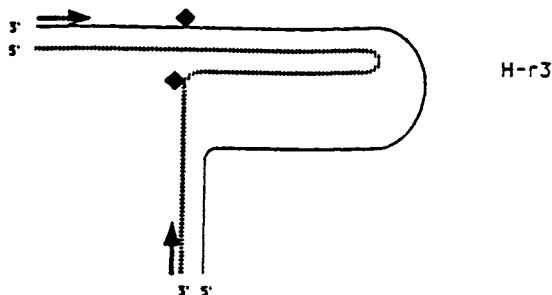
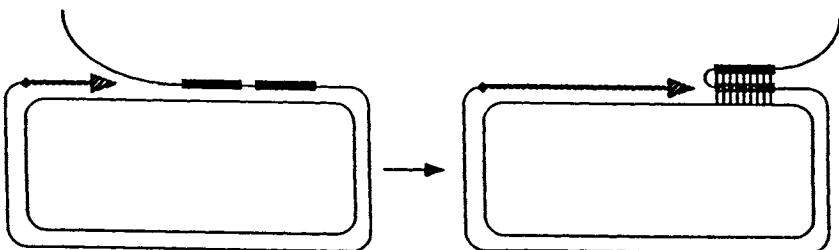
The unwound structure of the H-DNA may be appropriate for the binding of RNA polymerase (Figure 25C), which otherwise should itself (with the aid of aux-

iliary factors) locally unpair the DNA duplex to initiate transcription.⁴⁸⁴ Thus H-form could serve as the RNA polymerase entry point. Single-strand-binding proteins might transiently occupy a single-strand-repressing RNA polymerase binding when transcription is not required. The H-DNA structure appropriate for RNA polymerase entry could be formed under the influence of negative supercoiling (see above). Model studies in which RNA polymerases bound to locally unwound DNA templates and initiated⁵²⁰ and elongated transcription⁵²¹ support the RNA polymerase entry-point hypothesis.

In other cases a lack of correlation between the H-forming potential of the DNA sequence and the transcriptional efficiency of the promoter has been reported. A point mutation in the Py-Pu tract of the promoter of *Drosophila hsp26* gene destabilized the H-DNA *in vitro*, but had no effect on the level of gene expression.⁴⁷⁵ Replacement of one triplex-forming tract, $(GA)_n \cdot (TC)_n$, with another one, $(G)_n \cdot (C)_n$, reduced the transcription efficiency to the level characteristic of the promoter sequence lacking the essential $(GA)_n \cdot (TC)_n$ region. In this case, the H-DNA itself played no role in maintaining high transcriptional efficiency. A structure-function analysis of the c-Ki-ras promoter has shown that various point mutations in the Py-Pu tract do not affect its capability to form stable H-DNA *in vitro* and those destabilizing the triplex lead to a comparable drop in transcriptional activity relative to the original promoter.⁴⁹⁵ It was suggested that the Py-Pu sequences contained binding sites for specific proteins, and the mutations introduced might not affect H-forming potential, but reduce the protein binding affinity for the sequence.^{475,495} A similar conclusion has been drawn for the Py-Pu tract in the NCAM promoter.⁴⁹⁴ Among possible proteins that might bind to the Py-Pu sequences upstream of promoters are GAGA⁵²² and Sp1 transcription factors,⁴⁸⁵ NFκB,⁵²³ and NSEP proteins.⁵¹²

In addition to regulation at the initiation and termination stages, transcription in eukaryotes can be regulated at the elongation stage.⁵²⁴ The possible triplex role at the elongation stage may be illustrated by the experiments in which codon degeneracy was used to engineer a 38-bp H-forming sequence into the β -galactosidase gene of the pBluescriptIISK+ plasmid.⁴⁷⁸ An 80% lower expression of this β -galactosidase gene compared with another plasmid, where other codons that did not constitute the Py-Pu tract coded for the same amino acid sequence, was explained by the formation of H-DNA, which blocked the RNA polymerase progression along the DNA template.

In summary, the contradictory results on the importance of H-DNA in the regulation of transcription have been reported,^{475,495,519,525} and the mechanisms by which H-DNA-forming sequences influence gene expression are still far from being well understood.⁴⁰⁶ New information on mechanisms of H-DNA formation, and its stabilization and influence on conformations of adjacent regulatory sequences can advance our understanding of the potential involvement of H-DNA in biological processes.

A**B**

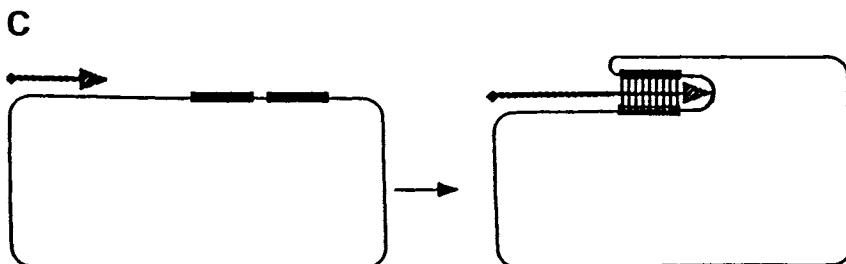


Figure 26. DNA polymerase inhibition by triplex DNA. (A) Inhibition by H*-DNA. The progression of a DNA polymerase along one strand of the double-stranded template may be hindered by the preexisting triple-stranded DNA structure. Depending on the H* DNA isomer, DNA polymerase advancing in the 3' to 5' direction faces the triplex either at the end or in the middle of the Py-Pu tract. (Reproduced from Mirkin and Frank-Kamenetskii.⁴⁰⁴) (B) Inhibition by H-DNA formation during strand displacement. When the new DNA chain is synthesized in a strand displacement reaction through the middle of the homopolymer template, the displaced strand containing a Py-Pu sequence may fold on itself downstream of the replication fork and form a block to replication. During the replication of open circular DNA under conditions favorable for both replication and H*-DNA formation, the T7 DNA polymerase stalled exactly in the middle of the Py-Pu tract when the purine-rich strand was displaced. (Reproduced from Samadashwily et al.⁵²⁹) (C) When a new DNA chain is synthesized through the middle of a single-strand homopolymer template (as occurs during synthesis of the lagging strand), the downstream half of the homopolymer sequence can fold back and form a triplex.⁵²⁷ This triplex serves as a trap for the DNA polymerase, which cannot continue DNA synthesis. (Reproduced from Mirkin and Frank-Kamenetskii.⁴⁰⁴).

2. Possible Regulation of Replication

A phenomenon similar to transcription elongation control was observed for DNA replication. The $(GA)_n \cdot (TC)_n$ tracts pause several DNA replication enzymes *in vitro* under conditions appropriate for triplex formation.^{526–528} Termination of DNA replication by the triplex-forming Py-Pu tracts has also been found in other studies^{262,479,529} (see, for a review, Mirkin and Frank-Kamenetskii).⁴⁰⁴ There are several different triplex structures that might be involved in the replication blockage.

The H*-DNA consisting of C-G-G and T-A-T triads could exist in different structural isomers dependent on the specific sequences designed.²⁶² The termination sites were located differently for specific H*-forming regions, but were mapped precisely by the chemical probing to the triplex-forming sequence. Figure 26A shows the sites where DNA polymerase movement may be hindered by the triple-stranded DNA structure that has been formed prior to DNA synthesis. Dependent on the H*-DNA isomer, DNA polymerase moving in the 3' to 5' direction along the template strand might stall either at the end or in the middle of the Py-Pu tract.

An active DNA polymerase itself may create a DNA structure that blocks polymerization. When the new DNA chain has been synthesized through the middle of the Py-Pu template, another portion of the Py-Pu sequence may fold back and form a triplex (Figure 26B).⁵²⁷ During the replication of single-stranded DNA, a DNA polymerase is trapped in the center of the homopyrimidine or homopurine tract.⁵²⁷⁻⁵²⁹ Upon double-stranded DNA replication, the DNA polymerase continuously synthesizes a strand complementary to the leading strand, and the other, lagging strand is unpaired for sufficient periods of time. The Py-Pu tract-containing lagging strand may fold on itself downstream of the replication fork. The resulting triplex presents one more type of replication block (Figure 26C). In accordance with this model, in experiments on open circular DNA, which cannot form a triplex before replication, T7 DNA polymerase stalled exactly in the middle of the Py-Pu tract when the purine-rich strand was displaced.⁵²⁹ Further evidence of replication termination by the induced triplex came from a mutational analysis. H*-DNA-destabilizing mutations relieved the DNA polymerase from the polymerization block, whereas the compensatory mutations restoring the H*-DNA-forming potential restored the replication block.

The models for DNA replication blockage by triplex structures were elaborated in experiments *in vitro*. Besides DNA polymerase, the actual replication fork contains a number of accessory proteins that may destroy the preformed triplex, prevent DNA single strands from folding onto their target duplex, etc. The single strand-binding protein easily restores replication by unwinding the intramolecular triplex, but is less effective in disrupting an intermolecular triplex.⁵³⁰ In *in vitro* experiments, the *E. coli* and the SV40 large T antigen DNA helicases were able to unwind model intermolecular triplexes⁵³¹ and intramolecular triplexes.⁵³² *In vivo* data on a role of Py-Pu tracts in replication are scarce. The (GA)₂₇-(TC)₂₇ tract-containing region of DNA located 2 kb from the integration site of the polyoma virus is a strong terminator of DNA replication in rat cells.⁵³³ The involvement of this Py-Pu tract in replication termination was confirmed by the DNA polymerase pausing when a corresponding fragment was cloned into SV40 DNA.⁵²⁶ Brinton et al.⁵³⁴ found that an unusual cluster of simple repeats, including a Z-DNA-forming region, (GC)₅(AC)₂₁, and a long Py-Pu tract with a potential for H-DNA formation, has a significant effect on replication of a plasmid shuttle vector. One copy of this cluster, when cloned on either side of SV40 origin of replication, reduced the amount of DNA replicated in COS cells up to twofold. Two copies on both sides of the origin reduced replicated DNA down to 5% of that in a vector without the cluster.

3. Possible Triplex-Mediated Chromosome Folding

The fact that the Py-Pu tracts are distributed over the whole length of genomic DNA^{496,497,535,536} has led to the suggestion that triplexes may promote chromosome compact packaging.^{537,538} A series of experiments was designed to show DNA condensation due to triplex formation between different DNA molecules or



Figure 27. Triplex DNA structures can form at telomeres. Telomeres have a single-strand end of G-rich sequences that can form a variety of triplex and quadruplex structures (see Section G). The single-stranded overhang (T_2G_4)₂ and DNA duplex of the same sequence can form a triplex in the presence of Mg^{2+} cations at physiological pH. Since the triple-stranded structure is not an appropriate substrate for many proteins, this triplex structure of telomeres may provide a plausible explanation for the *in vivo* resistance of chromosome ends against degradation and recombination. (Reproduced from Vesselkov et al.⁵⁴¹)

distant Py-Pu tracts of the same molecule.^{472,538} When a molecule of poly[d(Tm⁵C)] formed three-stranded structures with the Py-Pu tracts of several (GA)₄₅(TC)₄₅ insert-containing plasmids, this was seen in an electron microscope as the plasmid "rosettes."⁴⁷² In other experiments, linear plasmids containing single Py-Pu tracts at the ends produced linear dimers via triplex formation, and linear plasmids with Py-Pu tracts at both ends gave rise to quasicircular DNA molecules.⁵³⁸ In the above-described experiments, a triplex structure is formed between the Py-Pu duplex and one strand of another unwound PyPu tract. The action of topoisomerase or other nicking-closing activity is necessary to create the structure consisting of interwound strands, which belong to different long DNA molecules.

4. Structural Role at Chromosome Ends

Telomeres are structures that stabilize the ends of eukaryotic chromosomes. They usually present a very long repeating motif consisting of six to eight nucleotides with the general sequence $(T/A)_m G_n$.⁵³⁹ Single-stranded oligonucleotides modeling telomeric structures, consisting of the DNA duplex and the single-stranded overhang, form inter- and intramolecular quadruplexes in the presence of monovalent sodium and potassium cations.⁵⁴⁰

The study of a synthetic model of *Tetrahymena* chromosome telomeric terminus, consisting of the DNA duplex and the single-stranded overhang (T_2G_4)₂, showed that in the presence of Mg^{2+} and physiological pH the overhang folds back to form a triplex (Figure 27).⁵⁴¹ These authors suggested that the triplex structure of telomeres provides a plausible explanation for the *in vivo* resistance of chromosome ends against degradation and recombination.

5. Possible Role in Recombination

Specific types of triple-stranded structures have been suggested to mediate homologous DNA strand recombination in the presence of RecA protein.⁵⁴²⁻⁵⁴⁵

Such triple-stranded structures do not require the Py-Pu tracts, patterns of hydrogen bonding in them drastically differ from those in triplexes formed in the Py-Pu tracts, the axial spacing between base pairs is stretched to 5.1 Å, and homologous strands are aligned in parallel fashion.

The conventional intramolecular triplexes may also play roles in recombination. The DNA rearrangement in the immunoglobulin class from IgM to IgA, IgG, or IgE occurs in the highly repetitive regions of DNA (switch regions). For instance, the potentially H-DNA-forming sequence (AGGAG)₂₈ is located in the switch region of murine IgA.⁵⁴⁶ The unwound structure may provide a single strand to pair with a homologous region of a second chromosome-initiating recombination, which should result in the splicing of antibody genes from a number of individual gene segments. Non-B-DNA structures are possibly involved in an unequal sister chromatid exchange, in which part of the gene is duplicated on one chromosome and deleted from the other chromosome.⁵⁴⁷ The recombination region contains simple repeats of (TC)_n capable of forming H-DNA followed by stretches of the Z-forming (TG)_n sequence. Formation of dimer molecules in recombinant plasmids carrying the H-forming Py-Pu tracts occurs six times more often than in a control plasmid.⁵⁴⁷ In human cells homologous recombination between the plasmids containing H-DNA-forming sequences occurs three times more often compared to controls.⁵⁴⁸

Recombination can be transcriptionally induced between two direct repeats separated by the sequence containing the H*-forming Py-Pu tract (Figure 28).⁵⁵¹ The (dG)_n·(dC)_n-containing plasmid constructs allowed efficient recombination between homologous sequences of *lac* and *tac* promoter sequences separated by either 200- or 1000-bp regions. The recombination was RecA independent, the recombination rate being dependent on the length and orientation of (dG)_n·(dC)_n with respect to the gene. Under the active transcription conditions in *E. coli*, the plasmids formed C-G-G-type triplexes, as shown by chemical modification. The H*-DNA in this study was suggested to bring two remote sequences in close proximity to make recombination favorable.

The above described hypothetical models of H-DNA in genetic recombination may include a displacement mechanism (Figure 29A), an interaction of the H- and H*-forms (Figure 29B), and an H-form-induced duplex bend that brings homologous sequences into close proximity. Several results showing increased rates of plasmid dimerization, transcription-driven recombination between direct repeats, and the presence of H-forming tracts close to recombination points are indirect evidence in favor of the triplex playing a role in recombination. More experimental and theoretical considerations are needed to fully establish this role.

G. Control of Gene Expression

The formation of triple-stranded DNA complexes presents one of the most important examples of how DNA function can be influenced by changing its struc-

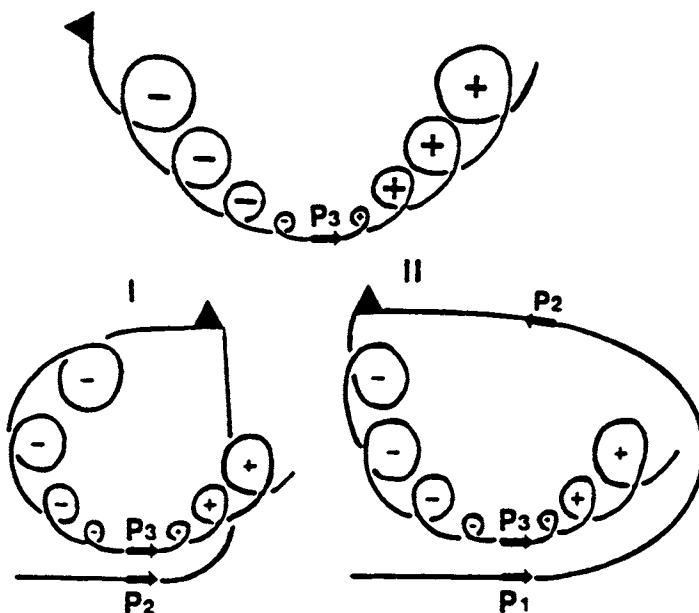


Figure 28. Transcription-stimulated genetic recombination mediated by triplex formation. Intramolecular triplex structures introduce bends into the DNA, which can influence the positioning of various DNA sequences. This model shows the triplex-induced approach of different homologous sequences in promoter regions (designated P1, P2, and P3). Once the promoter containing P3 is activated, the transcriptional process results in two supercoiled domains: positive supercoiling accumulates ahead of the transcribing RNA polymerase, whereas negative supercoiling accumulates behind RNA polymerase. Negative supercoiling-induced H*-DNA (designated by a filled triangle) significantly bends the DNA helix, bringing P2 and P3, which are separated by 200 bp, into close proximity to stimulate recombination. Similarly, P1 and P3, which are 1000 bp apart, may be brought together. (Reproduced from Kohwi and Panchenko.⁵⁴⁹)

ture. Intermolecular triplex formation at specific regulatory sites using specifically designed oligonucleotides may be used to inhibit deleterious expression of genes related to cancer, AIDS, and so forth. Triplex-based inhibition of gene expression is feasible because of a natural abundance of Py-Pu tracts in genes and their 5' flanking regions (see Table 6.1 in ref. 2).

Morgan and Wells⁵⁵⁰ showed that mRNA synthesis on a duplex polymer template is inhibited when the binding of a homopyrimidine third strand results in a triple-helical complex. During the past few years, triplex formation was shown to inhibit transcription and subsequent protein synthesis in a number of complicated

A. D-Loop Formation with an Unpaired Strand

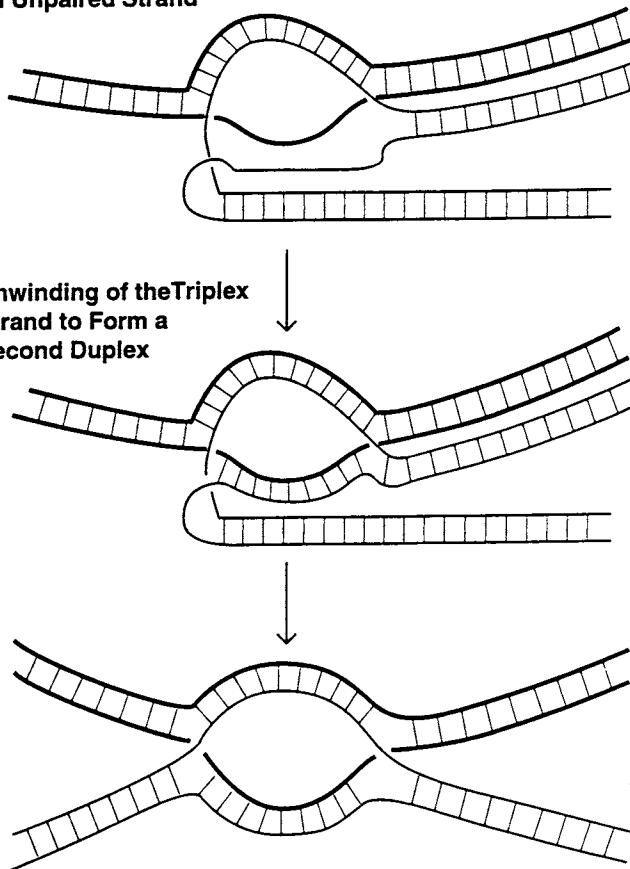
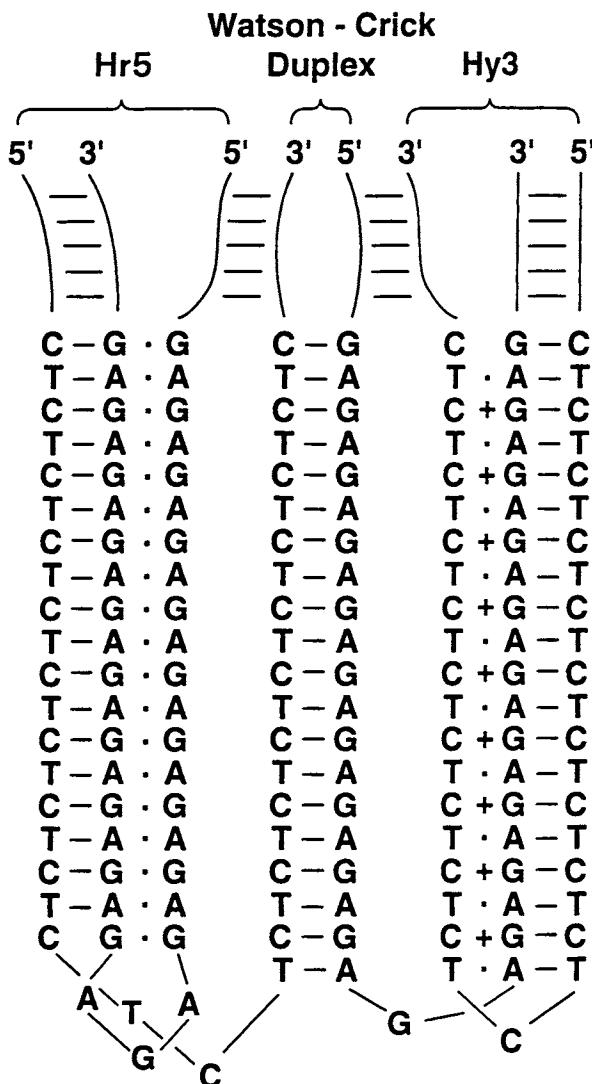


Figure 29. Hypothetical involvement of an intramolecular triplex in genetic recombination. (A) Possible mechanism of H-DNA-mediated genetic recombination. A free single strand of H-DNA pairs a complementary strand of a second homologous duplex (top). Following D-loop formation, the third strand of the triplex unpairs. Being complementary to the displaced D-loop strand, it initiates the formation of the Watson-Crick duplex (middle). Rotation of the bottom duplex from right to the left shows that a classical recombination intermediate has been formed (bottom). (Reproduced from Sinden, 1994.¹) (B) Recombination may involve two H-DNA structures. This model suggests the formation of different (H-r5 and H-y3; or H-r3 and H-y5) isomers in different DNA molecules containing the same Py-Pu sequence. Single strands of these H and H* forms are complementary and form a Watson-Crick duplex. The unwinding of the strands forming triple helices and their pairing in the Watson-Crick duplexes result in a classical recombination intermediate. (Reproduced from Sinden.¹)

B.



experimental cell-free systems. This inhibition is based on sufficiently high affinities of triplex-forming oligonucleotides for double-stranded DNA ($K_{diss} \approx 1-10$ nM) and the lifetimes of triple-stranded complexes on the order of several hours.^{431,432,437} These values approach those typical for many sequence-specific DNA-binding proteins, so the proteins involved in transcription cannot easily displace the Py-Pu tract-bound oligonucleotide. Specificity of triplex formation is

high: even one mismatch in a 15 nucleotide long triplex-forming oligonucleotide results in at least a 10-fold decrease in affinity compared to a perfect triplex.⁴¹⁴

Several mechanisms are relevant to the transcription inhibition by the triplex-forming oligonucleotides (Figure 30).^{403,410,551} The transcription machinery generally involves RNA polymerase and associated factors, activating proteins that bind upstream in the promoter region, and activators that bind to enhancer sequences at long distances from the RNA polymerase binding site and act via folding of the double helix. Triplex-forming oligonucleotides can (1) prevent long distance interactions between enhancer-bound proteins and RNA polymerase and associated factors by influencing the local bending ability of DNA; (2) eliminate binding of activators, acting either through long distance interactions or through factors bound adjacent to the RNA polymerase (not shown); (3) block tracking of transcription factors initially bound to a distant enhancer sequence and sliding toward the RNA polymerase machinery bound around the promoter site; (4) repress binding and interactions of basal protein factors associated with the RNA polymerase; (5) inhibit the RNA polymerase binding to the promoter site; (6) block initiation of transcription upon binding downstream of promoter but in contact with the transcription machinery; (7) inhibit elongation of transcription when bound to the Py-Pu site within the transcribed gene. Other mechanisms, such as the recruitment of inhibitory factors (e.g., proteins recognizing triple-stranded structure) or the alteration of chromatin assembly, may also be suggested. More details may be found in recent reviews.^{408,410}

Table 8 lists some examples of inhibition of transcription initiation or chain elongation by triplex-forming oligonucleotides. In addition to *in vitro* inhibition of gene expression, experiments have shown that the inhibiting activity is retained when the preformed DNA-triplex-forming oligonucleotide complex is introduced into cultured cells.^{552,553} Moreover, incubation of the cells in the presence of triplex-forming oligonucleotides resulted in the uptake of the oligonucleotides and a subsequent inhibition of gene expression and protein synthesis.⁵⁵⁴⁻⁵⁵⁸

In *in vitro* experiments, triplex-forming oligonucleotides, when used in excess to their target duplexes (mol oligonucleotide/mol template > 100), significantly inhibited (up to 90%) RNA polymerases.^{551,558,565} In *in vivo* experiments, 50% inhibition of mRNA synthesis or cell proliferation may be obtained at oligonucleotide concentrations in extracellular milieu of up to 100 µM.⁵⁵⁵⁻⁵⁵⁸ Inhibition of transcription depends on lifetimes of oligonucleotides, which in enzyme-rich cellular media are on the order of several dozen minutes^{452,566}; however, a decreased level of mRNA was noticed, even 7 days after removal of natural phosphodiester oligonucleotides from cell culture.⁵⁵⁷

In many cases the Py-Pu tracts are available in the genes encoding the key proteins in the pathogenesis of various diseases so that triplex methodology is applicable. Table 9 lists some diseases that may be treated with triplex-forming oligonucleotides and their analogs.⁴⁰⁷ The least complicated and most likely therapeutic applications of triplex-forming oligonucleotides will be as antiviral

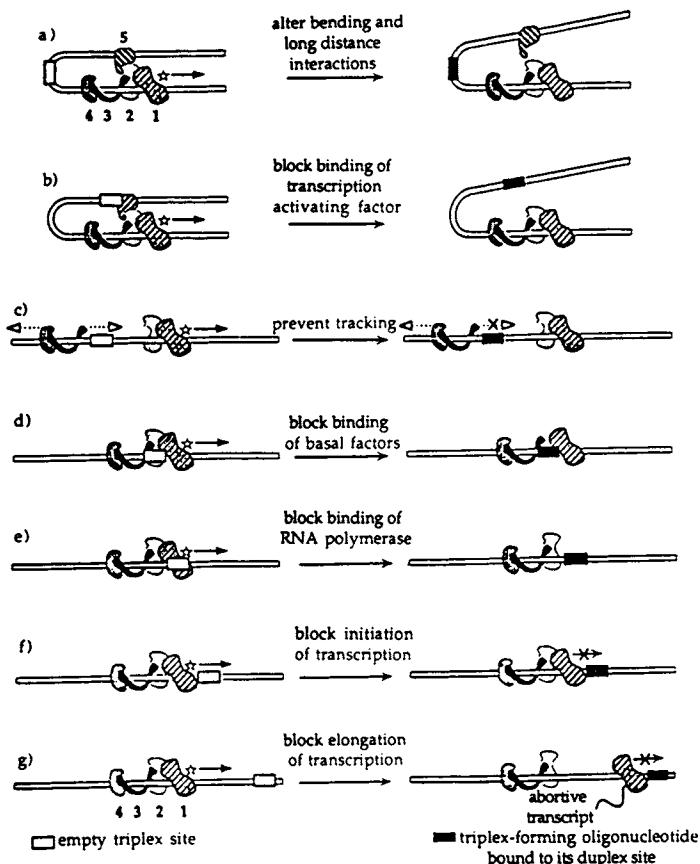


Figure 30. Hypothetical mechanisms by which triplex-forming oligonucleotides can inhibit transcription. The transcription process involves RNA polymerase and associated protein factors, activating factors that have binding sites in the promoter region, and activator proteins that bind to enhancer sequences at long distances upstream of the RNA polymerase binding site and require the double helix to fold. The star and arrow indicate the transcription start and the direction of transcription, respectively. An open box represents an unoccupied PyPu tract, whereas a filled box represents an intermolecular triplex formed on duplex DNA. Proteins: RNA polymerase (1), associated basal transcription factors (2), transcription activating factors (3,4,5). Triplex formation may result in (a) disruption of long-distance interactions between enhancer-bound proteins and the transcription complex; (b) blockage of activator binding at the distant site; (c) blockage of transcription factor sliding to the promoter sequence from the enhancer sequence; (d) inhibition of binding of basal transcription factors; (e) a physical barrier to RNA polymerase binding; (f) inhibition of transcriptional initiation upon binding downstream but in contact with transcription complex; (g) transcriptional inhibition at the elongation step when the Py-Pu tract is located within the transcribed region of the gene. (Reproduced from Thuong and Hélène.⁴⁰³)

Table 8. Inhibition of Transcription Via Intermolecular Triplex Formation

<i>Gene</i>	<i>Oligomer</i>	<i>Conditions</i>	<i>Reference</i>
Human c-myc	27-mer	<i>In vitro</i>	Cooney et al. ⁴¹⁶
Human c-myc	27-mer	HeLa cells (TFO uptake)	Postel et al. ⁵⁵⁵
Mouse IL2R	28-mer	Lymphocytes (TFO uptake)	Orson et al. ⁵⁵⁴
Mouse IL2R	15-mer-acridine	Tumor T cells HSB2 cells	Girgoriev et al. ⁵²³
	15-mer-psoralen	(plasmid-TFO electroporation)	Girgoriev et al. ⁵⁵²
<i>E. coli bla</i>	13-mer	<i>In vitro</i>	Duval-Valentin et al. ⁵⁵⁹
Human dihydrofolate reductase	19-mer	<i>In vitro</i>	Gee et al. ⁴⁸⁵
HIV-1	Various	<i>In vitro</i>	Ojwang et al. ⁵⁵⁷
	31 and 38-mers	Human MT4 and U937 cells (TFO uptake)	McShan et al. ⁵⁶⁰
T7 early promoter		<i>In vitro</i>	Ross et al. ⁴⁸⁶
Maize Adh1-GUS	Various lengths	Protoplasts (DNA and TFO co-transformation)	Lu and Ferl ⁵⁵³
Human platelet-derived growth factor A-chain	24-mer	<i>In vitro</i>	Wang et al. ⁵⁶¹
6-16 interferon-responsive element	21-mer	HeLa cells (liposome-mediated TFO delivery)	Roy ⁵⁵⁸
Progesterone-responsive gene	38-mer	Monkey kidney CV-1 cells (cholesterol-modified TFO uptake)	Ing et al. ⁵⁵⁶
Aldehyde dehydrogenase	21-mer	Human hepatoma cell Hep G2 (liposome-mediated TFO delivery)	Tu et al. ⁵⁶²
Rat $\alpha 1(I)$ collagen	30-mer	Rat cardiac fibroblasts (transfection) (liposome-mediated TFO delivery)	Kovacs et al. ⁵⁶³
Granulocyte-macrophage colony-stimulating factor	15-mer	Jurkat T cells (TFO uptake)	Kochetkova and Shannon ⁵⁶⁴

Table 9. Diseases that May Potentially Be Treated with Triplex-forming Oligonucleotides and Their Analogs

Virus-associated diseases

Adenovirus, herpes simplex viruses 1 and 2, herpes zoster, cytomegalovirus, Epstein-Barr virus, human papilloma virus, influenza A and B, parainfluenza, human T cell lymphotropic virus, human immunodeficiency virus, hepatitis A and B

Oncologic diseases

Lymphoma, leukemia, melanoma, osteosarcoma, carcinoma (of colon, prostate, kidney, bladder, breast)

Other

Psoriasis, drug resistance, allergy, inflammation

agents. For many pathogenic viruses, the protein sequences and functions are relatively well understood. Many essential viral proteins whose genes should be suppressed have no human cellular analogs, which reduces the risk of undesirable suppression of normally functioning human genes. The prospects for the treatment of cancer diseases with triplex-forming oligonucleotides are less clear. Although oncogenes were suggested in the development and progression of several human tumors, the understanding of underlying processes is not clear enough. Moreover, the difference between an oncogene and its normal cellular counterpart may be in only one base pair that presents a risk of unspecific inhibition of normal cellular genes.

There are other important problems in the triplex-mediated inhibition of gene expression.^{2,410,567}

1. The rate of triplex formation is slow, therefore, for oligonucleotides competing with protein factors for binding to a specific sequence on DNA, kinetic phenomena might become a limiting factor.
2. The range of limitations imposed by nonspecific oligonucleotide binding to numerous target sites (e.g., transfer RNA, small nuclear RNA, accessible single-stranded regions in ribosomal RNA, aminoacyl-tRNA-synthetases, false binding to the initiation start, etc.) remains to be elucidated.
3. Triplex-stabilizing low pH or elevated concentrations of divalent metal cations and polyamines are not available in cells. Therefore, cytosine analogs, which are not protonated at neutral pH, and stably binding purine-rich oligonucleotides containing T and G nucleotides are being developed (see ref. 2 for a review). Unfortunately, cytosine-rich pyrimidine oligonucleotides may form protonated hairpin structures, reducing the oligonucleotide amount available for triplex formation,⁵⁶⁸ and T- and G-containing oligonucleotides may aggregate at physiological ionic strength by forming guanosine quartets.^{465,569}

4. Oligonucleotides designed to form triplexes may act by interrupting other cellular processes (e.g., the cascade of events required for interferon-mediated induction of several genes, including the targeted one).^{570,571} Thus an interpretation of oligonucleotide effects *in vivo* requires some caution.
5. In some cases, short oligonucleotides may not bind strongly enough to stop the enzymatic machinery. Cross-linking or intercalating agents covalently linked to oligonucleotides (Table 6) might be used to stabilize the triplexes. However, the issues of a cross-link repair and inhibition of DNA replication become important.⁴¹⁰
6. Delivery of polyanionic, hydrophilic compounds to their targets through the lipid membrane must be efficient enough.⁵⁷² Since negatively charged oligonucleotides poorly penetrate hydrophobic cellular membranes, several techniques of oligonucleotide delivery into the cell via liposome-⁵⁷³⁻⁵⁷⁵ and receptor-mediated endocytosis,⁵⁷⁶ directly in the hydrophobized form⁵⁵⁶ and as conjugates with polylysine,⁵⁷⁷ have been developed. The oligonucleotides of interest can be generated at their action sites from a vector containing promoter, capping, and termination sequences of the human small nuclear U6 gene, surrounding a synthetic sequence to be synthesized.
7. When delivered into the cell, the full-length oligonucleotides may be stable enough to persist for several hours.⁵⁵⁴ A number of other analogs have been designed to improve the triplex-forming ability and metabolic stability (see ref. 2 for a review). There is growing evidence that oligonucleotides enter cells by endocytosis into lysosomes and are inefficiently delivered to the nucleus,⁴¹⁰ which requires the modification of existing or development of new delivery strategies.

In summary, any disease caused by the expression of a gene can be treated at various stages of the cellular processes. Many traditional drugs are targeted against the functional proteins, and their design requires extensive information about structure-activity relationships for these drugs. The oligonucleotide-based antisense approach is directed against the formation on the mRNA template of the secondary products, proteins. However, a continuous gene expression may supply new mRNA molecules. The triplex-based antigene strategy exploits the possibility of binding oligonucleotides to DNA inside genes or regulatory regions to hamper mRNA synthesis on the DNA template. It is directed against the primary process of gene expression and, therefore, may be more efficient. Another advantage is that it avoids the necessity of determining the various specific mechanisms of drug-protein interaction. In a practical sense, several important problems must be solved before triplex-mediated gene therapy will become a reality.

VII. MISCELLANEOUS ALTERNATIVE CONFORMATIONS OF DNA

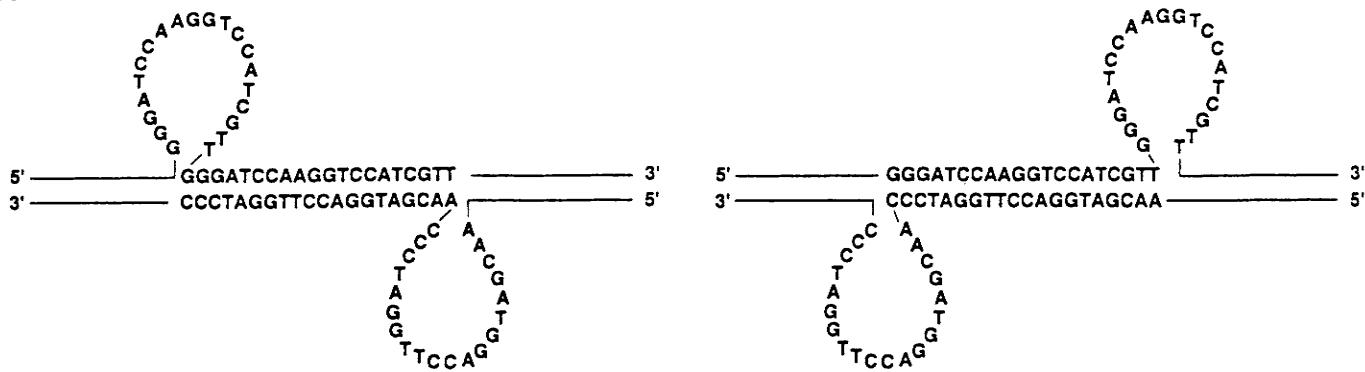
A. Slipped-Strand DNA

Slipped-strand DNA (S-DNA) structures can form in regions with direct repeat symmetry. Regions of DNA containing long tracts of repeating mono-, di-, tri-, and tetranucleotides have multiple opportunities for the formation of hydrogen bonds in an out-of-register or “slipped” fashion. To form a slipped-strand structure, a section (or all) of the repeating duplex must unwind to allow one region of the direct repeat to form a Watson–Crick base pair strand with another region of the repetitive sequence. Figure 31 shows two possible isomers of S-DNA. One isomer has loops composed of the 5' direct repeat in both strands, and the other has loops composed of the 3' direct repeats in both strands. Because this structure results in the unwinding of the DNA double helix, DNA supercoils would be lost upon the formation of S-DNA in supercoiled DNA.

Historically, S-DNA structures have been suggested to exist in eukaryotic DNA within regions of direct repeat symmetry sensitive to S1 nuclease.^{578–580} However, these sequences also contained mirror repeat symmetry, and intramolecular triplex structures may form at these sites. Recent evidence for S-DNA has been presented by Pearson and Sinden.⁵⁸¹ By melting and reannealing DNA, they detected novel structures formed within the (CTG)·(CAG) and (CGG)·(CCG) triplet repeat tracts associated with myotonic dystrophy and fragile X syndrome, respectively. The reannealing-induced CTG- or CGG-containing DNA structures have the following properties: (1) the novel structures are formed from complementary strands; (2) the structures are formed from complementary strands of equal length; (3) the alternative DNA structure occurs within the repeat tract; (4) linear duplex DNA flanks the alternative structure; (5) formation of the alternative structure does not require superhelical tension; (6) the alternative structures are remarkably stable in linear DNA under physiological conditions; (7) the alternative structures possess single-strand character, which is expected of DNA in two loops. Moreover, the CAG strand is more susceptible to single-strand nuclease digestion than is the CTG strand, consistent with reports of a greater stability of hairpin structures in the CTG strand (see Pearson and Sinden⁵⁸¹ for a discussion). The results are consistent with the formation of S-DNA structures (Figure 31).

Biologically, S-DNA is very important in spontaneous frameshift mutagenesis. It is known that spontaneous deletion or addition mutations can occur within runs of a single base. In 1966, Streisinger et al. proposed a model to explain frameshift mutations within runs of a single base.⁵⁸² Since the genetic code is read as triplets, adding or deleting a single base shifts the reading frame of all bases downstream of the mutation. This will result in a mRNA that encodes amino acids that are different from those present in the wild-type protein downstream of the frameshift mutation.

A



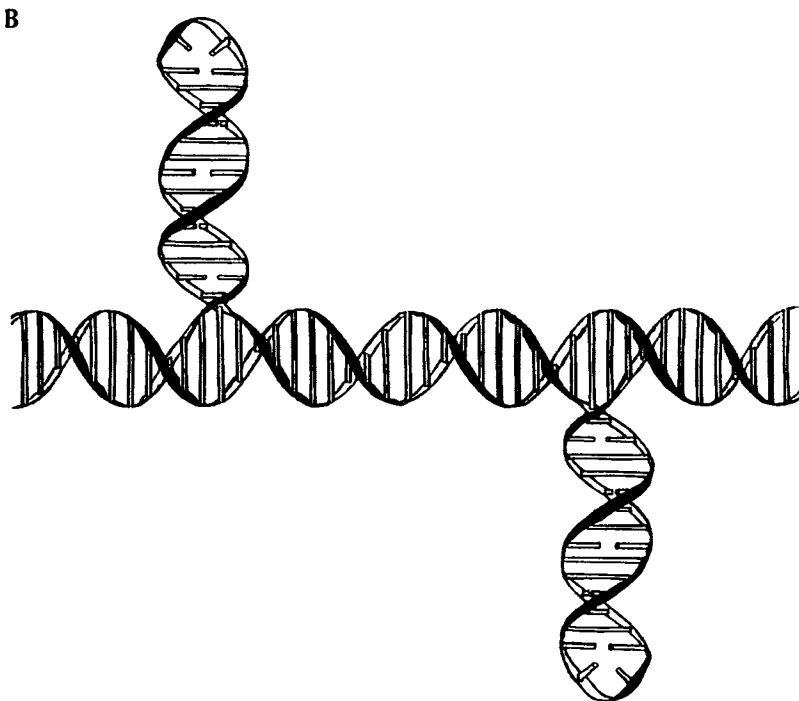


Figure 31. Slipped-strand DNA. Slipped mispaired DNA can form within two direct repeat sequences when they pair in a misaligned fashion. (A) Two 20-bp direct repeats can form two different slipped mispaired isomers. In one isomer (left), the second copy of the direct repeat in the top strand is base paired with the first copy of the direct repeat on the bottom strand. The other structure (right) shows an isomer in which the first copy of direct repeat in the top strand is base paired with the second copy of the direct repeat in the bottom strand. (B) S-DNA formed from $(CTG)_n \cdot (CAG)_n$ and $(CCG)_n \cdot (CCG)_n$ triplet repeats. S-DNA structures are believed to form within these triplet repeat sequences following denaturation and renaturation.⁵⁸¹ The looped-out single strands can fold into hairpin structures stabilized by two G-C base pairs flanking an A-A or T-T mispair in the opposite strands.

B. DNA Unwinding Elements

DNA unwinding elements or DUEs have been identified in both prokaryotic and eukaryotic DNA sequences. DUEs are A + T-rich regions of DNA that are commonly associated with replication origins and chromosomal matrix attachment sites. DUEs are A + T-rich sequences ranging in size from 30 to >100 bp in length. As a class of sequence elements, DUEs have little sequence homology in that there is no consensus sequence. The only similarity is that the sequences are A + T-rich.

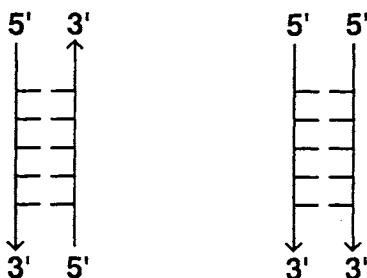
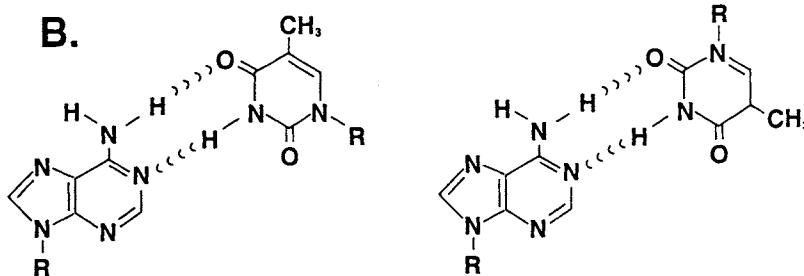
In the presence of DNA supercoiling, unwinding of the double helix occurs first at A + T regions. *In vitro* DUEs unwind, and this unwound state is maintained in a stable fashion in the presence of negative DNA supercoiling. This unwound state can be detected by the sensitivity of the DUEs to digestion by single-stranded nucleases. Shefflin and Kowalski examined the pattern of cutting supercoiled DNA with mung bean nuclease.⁵⁸³ The susceptibility of the DUEs to nuclease digestion is a function of the Mg²⁺ concentration. In buffer containing no Mg²⁺, DUEs melt and are sensitive to digestion, whereas in the presence of Mg²⁺, the A + T-rich DUE region remained double stranded. Thus the ability of these regions of DNA to unwind *in vivo* may be controlled by the level of unrestrained supercoiling and the local ionic environment in cells.

DUEs are required for the initiation of DNA replication at certain origins, as discussed above in Section IV^{202,205}. A correlation exists between DNA unwinding and the proficiency of the DUE as a replication origin, as determined from certain yeast sequences. Progressive deletion of the DUE decreased function as an origin of DNA replication until, when melting no longer occurred, the region no longer functioned as a replication origin. These results suggest that unwinding at an A + T-rich DNA region near the origin of DNA replication is required for the initiation of DNA replication. DNA unwinding is also required at the *E. coli* origin of replication. DnaA proteins first bind to DnaA boxes at the *E. coli* origin of replication in supercoiled DNA and organize DNA into a tight loop. Three DnaB boxes, which are A + T-rich direct repeat sequences (effectively DUE elements) and then unwind. Following binding of DnaB and DnaC, the DnaB helicase begins unwinding the DNA double helix, forming a bubble in which DnaG, an RNA primase, binds. Subsequently, DNA polymerase holoenzyme binds and replication elongation ensues.

C. Parallel-Stranded DNA

Parallel-stranded DNA contains the two single strands in an orientation opposite that of typical Watson–Crick helix orientation, which is antiparallel.^{584,585} The orientation of the two strands in parallel DNA requires that the bases be paired in a reverse Watson–Crick fashion. In the reverse Watson–Crick orientation, the glycosidic bonds and ribose sugars are extended in a *trans* position, compared with the *cis* configuration found in B-DNA (Figure 32). In a reverse Watson–Crick base pair, the thymines are hydrogen-bonded through the O2 and N3 positions, whereas the N3 and O4 positions are hydrogen bonded in a normal Watson–Crick base pair. Remarkably, the stability of parallel stranded DNA is only slightly lower than that of a corresponding parallel B-DNA helix.

There are no known examples of naturally occurring parallel DNA formed by A-T base pairs. However, G + C-rich regions in the chromosome and sequences found at telomeres can form duplex and quadruplex structures in which the strands are organized in a parallel fashion. Certain RNA sequences may also adopt a parallel helix.

A. **Anti Parallel** **Parallel**

B.


Watson-Crick A-T Base Pair

Reverse Watson-Crick A-T Base Pair

C.


Figure 32. Parallel-stranded DNA. (A) Antiparallel and parallel organization of complementary strands. Typical B-form, A-form, or Z-form DNA contains an antiparallel orientation of the complementary strands, with 5'→3' polarity in one strand and 3'→5' polarity in the opposite strand. (B) Parallel DNA requires the formation of reverse Watson-Crick base pairs in which one base is oriented 180° with respect to the Watson-Crick base pair orientation. (C) The A + T-rich sequence shown can form a parallel helix.⁵⁹⁶ The reverse Watson-Crick base pairs are indicated by an open circle (o).

D. Four-Stranded DNA

In recent years it has come to be realized that there are many ways four strands of DNA can be held together by various hydrogen bonding schemes forming quadruplex DNA. A G + C-rich region of DNA from the immunoglobulin heavy chain

switch region was shown to exist as a four-stranded, G-quartet structure in which all strands are parallel.⁵⁸⁶ On incubation of single-stranded oligonucleotides containing 5'-GGGGAGCTGGG-3', a higher molecular weight structure formed, as detected by electrophoresis on polyacrylamide gels. A Hoogsteen base pairing scheme involving four DNA strands organized in a parallel configuration was suggested from analysis of the chemical reactivity of the complex to dimethylsulfate. This G-quartet DNA structure may be important biologically. For example, a quadruplex might hold the four chromosomes together at meiosis.⁵⁸⁶

Quadruplex DNA can also form in DNA tracts found at telomeres. Telomeres are composed of purine-rich repetitive sequences (for example, (G₄T₂)_n or (G₄T₄)_n). A multitude of alternative pairing schemes are possible with these telomere repeats (for example, see ref. 587). The *Tetrahymena* telomere DNA sequence (G₄T₂) can form a quadruplex consisting of a planar array of guanines in the *anti* configuration held together by Hoogsteen base pairs.⁵⁸⁸ Depending on how these sequences come together, association can occur in a parallel or antiparallel orientation. A parallel orientation can occur from the interaction of four single strands (Figure 33). If a single strand first folds into a hairpin in which guanines are Hoogsteen base paired, the two strands are antiparallel. Two of these foldback structures can interact in four-stranded structures, as discussed below.

The formation of two different isomers of the *Oxytricha* telomeric DNA sequence (G₄T₄G₄) have been reported.^{589,590} The individual G₄T₄G₄ molecules can form hairpins, and two hairpins can hydrogen bond in two different ways. The organization with alternating strands arranged in an antiparallel orientation was found in an X-ray crystal structure by Kang et al.⁵⁸⁹ Alternate nucleotides in a quartet exist in the *anti* and *syn* conformations. The pattern and direction of hydrogen bonding are reversed in adjacent quartets (Figure 34). Smith and Feigon, using NMR analysis, reported a very different quadruplex organization of the same sequence: a four-stranded structure that does not involve hairpins.⁵⁹⁰ In this structure, the T₄ loops are on opposite ends of the quartet spanning the diagonals of the quartet at right angles to each other. In this structure, the glycosidic bond angles are *syn-syn-anti-anti* around the quartet, whereas the pattern in adjacent quartets is *anti-anti-syn-syn* (Figure 34). Like the double hairpin structure, the direction of hydrogen bonding also alternates between adjacent quartets.

The very different quartet structures identified for the same *Oxytricha* telomere sequence (G₄T₄G₄) by Smith and Feigon⁵⁹⁰ and Kang and colleagues⁵⁸⁹ demonstrate the variability inherent in the formation of quadruplex structures from poly(dG) sequences. Presumably the variable structures represent different local ionic and buffer conditions present during crystallization or NMR analysis. The structures that form are sensitive to monovalent cations.^{587,591,592} A single K⁺ ion may bind tightly within the center of the quartet. If quartet structures form in living cells, the structural isomer that forms, as well as their stability, might be controlled by variations in intracellular K⁺ concentration.

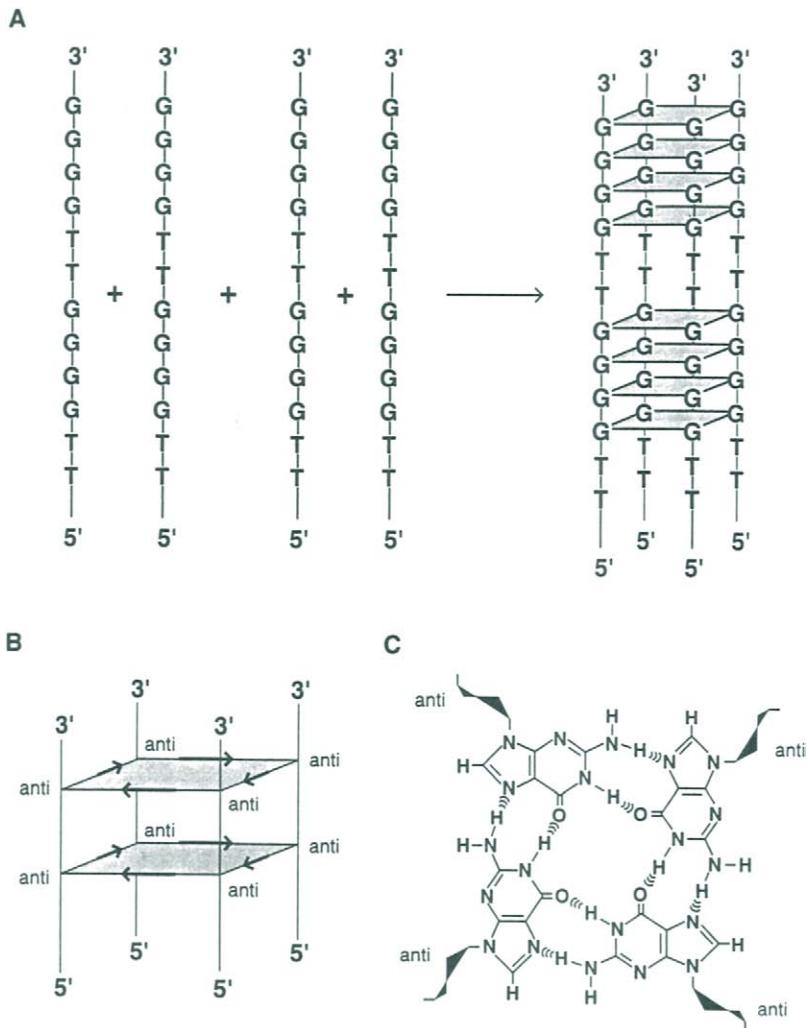


Figure 33. A G-quartet quadruplex DNA structure. (A) Four single strand tracts containing (G_4T_2) can form a tetraplex structure with all four strands in a parallel orientation. (B and C) The four G's, with all glycosidic bonds in the *anti* conformation, are hydrogen bonded through Hoogsteen base pairs. Moreover, the direction of hydrogen bonds are the same in each G quartet. The arrows (in B) indicate the direction of the hydrogen bonding from the hydrogen donor (H) to the acceptor (O or N). The direction is similar in successive base quartets.

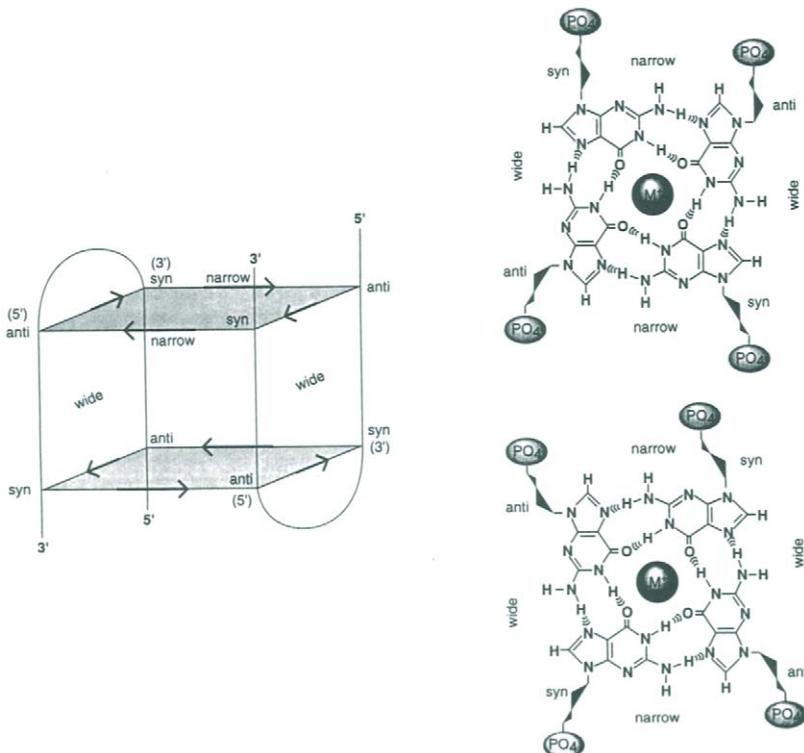
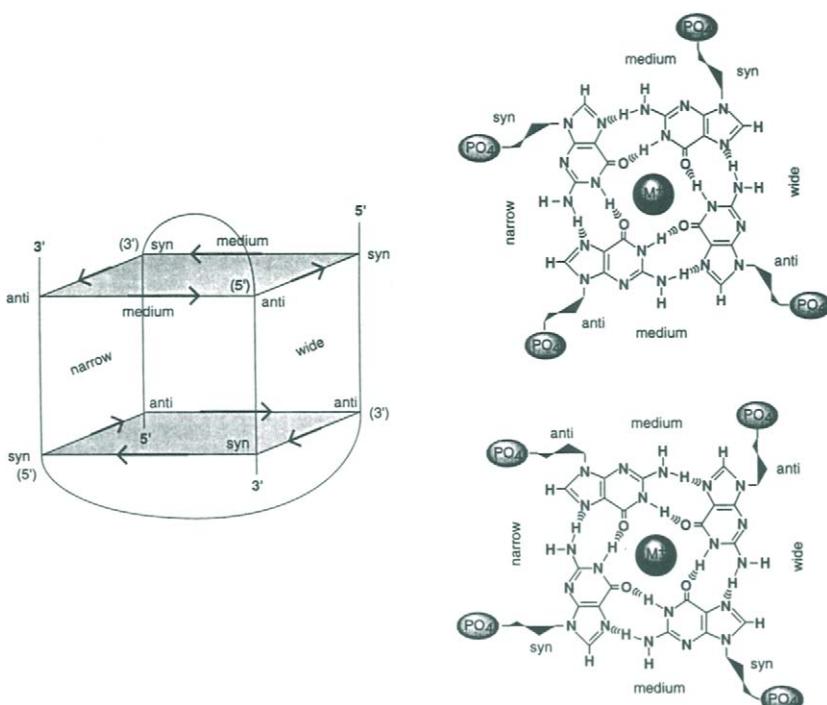
A

Figure 34. Different four-stranded DNA structures from a telomere sequence. Many different structural isomers can form from sequences found at telomeres. (A) A single strand containing $(G_4T_4)_n$ can fold into a hairpin that will be an antiparallel duplex. Two hairpins can associate in four different ways, with the hairpin loops at the same or opposite ends of the quadruplex. In each orientation the polarity of the DNA strands can be 5'-3'-5'-3' or 5'-5'-3'-3'. The *Oxytricha* telomere sequence G₄T₄G₄ crystallized into an antiparallel quadruplex, with the hairpin loops at opposite ends of the quartet and with 5'-3'-5'-3' polarity.⁵⁸⁹ Within a G quartet the glycosidic bonds alternate anti-syn-anti-syn. Moreover, the direction of hydrogen bonding alternates in adjacent quartets. (B) (G₄T₄)_n telomere sequences can also form a different structure formed when the two single strands fold together, forming the diagonals of a quartet with the loops at right angles to each other. There are no intrastrand hydrogen bonds, rather guanines from one strand hydrogen bond to two guanines from the second strand. There are two different isomers that can form, depending on the ways the two single strands fold. The structure shown was identified in solution for G₄T₄G₄ by NMR spectroscopy.⁵⁹⁰ In this orientation the glycosidic bonds are anti-anti-syn-syn in one quartet and syn-syn-anti-anti in an adjacent quartet. The direction of the hydrogen bonds is reversed between adjacent quartets.



E. Higher Order Pu·Py Structures

Long regions of $(dG)_n \cdot (dC)_n$ can form a bimolecular triplex from two smaller intramolecular triplex regions when the single strand loop from the first intramolecular triplex interacts to form the third strand of an adjacent intramolecular triplex.^{429,461} When a homopurine-homopyrimidine region gets very long, there are many different opportunities for smaller triple helices to form. This is especially pronounced in very simple repetitive sequence elements such as poly(dG)·poly(dC) or polyd(G-A)·polyd(T-C). Glover and Pulleyblank⁵⁹³ and Shimizu et al.⁵⁹⁴ demonstrated considerable structural diversity in long simple polypurine-polypyrimidine mirror repeats, as is evident from multiple bands indicative of different DNA secondary transitions on two dimensional gels.

REFERENCES

1. Sinden, R.R. *DNA Structure and Function*. Academic Press, San Diego, 1994.
2. Soyfer, V.N.; Potaman, V.N. *Triple-Helical Nucleic Acids*. Springer, New York, 1995.
3. Seeman, N.C.; Rosenberg, J.M.; Suddath, F.L.; Kim, J.J.P.; Rich, A. RNA double helical fragment at atomic resolution. I. The crystal and molecular structure of sodium adenylyl-3',5'-uridine hexahydrate. *J. Mol. Biol.* **1976**, *104*, 109–144.

4. Rosenberg, J.M.; Seeman, N.C.; Day, R.O.; Rich, A. RNA double helical fragment at the atomic resolution. II. The crystal structure of sodium guanylyl-3',5'-cytidine nonahydrate. *J. Mol. Biol.* **1976**, *104*, 145–167.
5. Frank-Kamenetskii, M.D. Simplification of the empirical relationship between melting temperature of DNA, its GC content and the concentration of sodium ions in solution. *Biopolymers* **1971**, *10*, 2623–2624.
6. Hoogsteen, K. The crystal and molecular structure of a hydrogen-bonded complex between 1-methylthymine and 9-methyladenine. *Acta Crystallogr.* **1963**, *16*, 907–916.
7. Langridge, R.; Wilson, H.R.; Hooper, C.W.; Wilkins, M.H.F.; Hamilton, L.D. The molecular configuration of deoxyribonucleic acid. I. X-ray diffraction study of a crystalline form of the lithium salt. *J. Mol. Biol.* **1960**, *2*, 19–37.
8. Langridge, R.; Marvin, D.A.; Seeds, W.E.; Wilson, H.R.; Hamilton, L.D. The molecular configuration of deoxyribonucleic acid II. Molecular models and their Fourier transforms. *J. Mol. Biol.* **1960**, *2*, 38–64.
9. Fuller, W.; Wilkins, M.H.F.; Wilson, H.R.; Hamilton, L.D.; Arnott, S. The molecular configuration of deoxyribonucleic acid. IV. X-ray diffraction study of the A form. *J. Mol. Biol.* **1965**, *12*, 60–80.
10. Fairall, L.; Martin, S.; Rhodes, D. The DNA binding site of the *Xenopus* transcription factor IIIA has a non-B-form structure. *EMBO J.* **1989**, *8*, 1809–1817.
11. Dickerson, R.E.; Drew, H.R. Structure of a B-DNA dodecamer II. Influence of base sequence on helix structure. *J. Mol. Biol.* **1981**, *149*, 761–786.
12. Dickerson, R.E. Base sequence and helix structure variation in B and A DNA. *J. Mol. Biol.* **1983**, *166*, 419–441.
13. Arnott, S.; Chandrasekaran, R.; Hall, I.H.; Puigjaner, L.C. Heteronomous DNA. *Nucleic Acids Res.* **1983**, *11*, 4141–4156.
14. Chastain, P.D.; Eichler, E.E.; Kang, S.; Nelson, D.L.; Levene, S.D.; Sinden, R.R. Anomalous rapid electrophoretic mobility of DNA containing triplet repeats associated with human disease genes. *Biochemistry* **1995**, *34*, 16125–16131.
- 14a. Chastain, P.D.; Sinden, R.R. CTG repeats associated with human genetic disease are inherently flexible. *J. Mol. Biol.* **1997**, in press.
15. Hagerman, P.J. Flexibility of DNA. *Annu. Rev. Biophys. Biophys. Chem.* **1988**, *17*, 265–286.
16. Bednar, J.; Furrer, P.; Katritch, V.; Stasiak, A.Z.; Dubochet, J.; Revet, B. Determination of DNA persistence length by cryoelectron microscopy—separation of the static and dynamic contributions to the apparent persistence length of DNA. *J. Mol. Biol.* **1995**, *254*, 579–591.
17. Marini, J.C.; Levene, S.D.; Crothers, D.M.; Englund, P.T. Bent helical structure in kinetoplast DNA. *Proc. Natl. Acad. Sci. USA* **1982**, *79*, 7664–7668.
18. Marini, J.C.; Levene, S.D.; Crothers, D.M.; Englund, P.T. Bent helical structure in kinetoplast DNA. *Proc. Natl. Acad. Sci. USA* **1983**, *80*, 7678.
19. Bustamante, C.; Gurrieri, S.; Smith, S.B. Towards a molecular description of pulsed-field gel-electrophoresis. *Trends Biotechnol.* **1993**, *11*, 23–30.
20. Calladine, C.R.; Drew, H.R.; McCall, M.J. The intrinsic curvature of DNA in solution. *J. Mol. Biol.* **1988**, *201*, 127–137.
21. Maroun, R.C.; Olson, W.K. Base sequence effects in double-helical DNA. II. Configurational statistics of rod-like chains. *Biopolymers* **1988**, *27*, 561.
22. Travers, A.A. DNA conformation and protein binding. *Annu. Rev. Biochem.* **1995**, *58*, 427–452.
23. Wu, H.M.; Crothers, D.M. The locus of sequence-directed and protein-induced DNA bending. *Nature* **1984**, *308*, 509–513.
24. Koo, H.S.; Wu, H.M.; Crothers, D. DNA bending at adenine-thymine tracts. *Nature* **1986**, *320*, 501–506.
25. Hagerman, P.J. Sequence dependence of the curvature of DNA: A test of the phasing hypothesis. *Biochemistry* **1985**, *24*, 7033–7037.

26. Brukner, I.; Dlakic, M.; Savic, A.; Susic, S.; Pongor, S.; Suck, D. Evidence for opposite groove-directed curvature of GGGCCC and AAAA sequence elements. *Nucleic Acids Res.* **1993**, *21*, 1025–1029.
27. Brukner, I.; Susic, S.; Dlakic, M. Physiological concentration of magnesium ions induces a strong macroscopic curvature in GGGCCC-containing DNA. *J. Mol. Biol.* **1994**, *236*, 26–32.
28. Trifonov, E.N.; Sussman, J.L. The pitch of chromatin DNA is reflected in its nucleotide sequence. *Proc. Natl. Acad. Sci. USA* **1980**, *77*, 3816–3820.
29. Ulanovsky, L.; Bodner, M.; Trifonov, E.N.; Choder, M. Curved DNA: Design, synthesis, and circularization. *Proc. Natl. Acad. Sci. USA* **1986**, *83*, 862–866.
30. Trifonov, E.N.; Tan, R.K.Z.; Harvey, S.C. Static persistence length of DNA. In: *Structure and Expression* (Sarma, M.H.; Sarma, R.H.; eds.). Adenine Press, Albany, NY, 1988, pp. 243–254.
31. Koo, H.S.; Drak, J.; Rice, J.A.; Crothers, D.M. Determination of the extent of DNA bending by an adenine-thymine tract. *Biochemistry* **1990**, *29*, 4227–4234.
32. Selzing, E.; Wells, R.D.; Alden, C.J.; Arnott, S. Bent DNA: visualization of a base-paired and stacked A-B conformational junction. *J. Biol. Chem.* **1979**, *254*, 5417–5422.
33. Goodsell, D.S.; Kopka, M.L.; Cascio, D.; Dickerson, R.E. Crystal structure of CATGGCCATG and its implications for A-tract bending models. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 2930–2934.
34. Brahms, S.; Brahms, J.G. DNA with adenine tracts contains poly(dA)-poly(dT) conformational features in solution. *Nucleic Acids Res.* **1990**, *18*, 1559–1564.
35. Nelson, H.C.M.; Finch, J.T.; Luisi, B.F.; Klug, A. The structure of an oligo(dA)-oligo(dT) tract and its biological implications. *Nature* **1987**, *330*, 221–226.
36. Coll, M.; Frederick, C.A.; Wang, A.H.; Rich, A. A bifurcated hydrogen-bonded conformation in the d(A-T) base pairs of the DNA dodecamer d(CGCAAATTTCG) and its complex with distamycin. *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 8385–8389.
37. Peck, L.; Wang, J.C. Sequence dependence of the helical repeat of DNA in solution. *Nature* **1981**, *292*, 375.
38. Rhodes, D.; Klug, A. Sequence-dependent helical periodicity of DNA. *Nature* **1981**, *292*, 378–380.
39. Burkhoff, A.M.; Tullius, T.D. Structural details of an adenine tract that does not cause DNA to bend. *Nature* **1988**, *331*, 455–457.
40. Burkhoff, A.M.; Tullius, T.D. The unusual conformation adopted by the adenine tracts in kinetoplast DNA. *Cell* **1987**, *48*, 935–943.
41. McCarthy, J.G.; Frederick, C.A.; Nicolas, A. A structural analysis of the bent kinetoplast DNA from *fasciculata* by high resolution chemical probing. *Nucleic Acids Res.* **1993**, *21*, 3309–3317.
42. Young, M.A.; Ravishankar, G.; Beveridge, D.L.; Berman, H.M. Analysis of local helix bending in crystal-structures of DNA oligonucleotides and DNA-protein complexes. *Biophys. J.* **1994**, *68*, 2454–2468.
43. Nadeau, J.G.; Crothers, D.M. Structural basis for DNA bending. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 2622–2626.
44. Chen, S.M.; Leupin, M.; Chazin, W.J. Conformational studies of the duplex d-(Caaaaatttcc)-d-(ggaaattttgg) containing a (dA)₅ tract using two-dimensional ¹H-n.m.r. spectroscopy. *Int. J. Biol. Macromol.* **1992**, *14*, 57–63.
45. Leroy, J.L.; Charretier, E.; Kochyan, M.; Gueron, M. Evidence from base-pair kinetics for two types of adenine tract structures in solution: their relation to DNA curvature. *Biochemistry* **1988**, *27*, 8894–8898.
46. Diekmann, S. Temperature and salt dependence of gel migration anomaly of curved DNA fragments. *Nucleic Acids Res.* **1987**, *15*, 247–265.
47. Chan, S.S.; Breslauer, K.J.; Hogan, M.E.; Kessler, D.J.; Austin, R.H.; Ojemann, J. et al. Physical studies of DNA premelting equilibria in duplexes with and without homo dA-dT tracts: correlations with DNA bending. *Biochemistry* **1990**, *29*, 6161–6171.
48. Chan, S.S.; Breslauer, K.J.; Austin, R.H.; Hogan, M.E. Thermodynamics and premelting conformational changes of phased (dA)₅ tracts. *Biochemistry* **1993**, *32*, 11776–11784.

49. de Souza, O.N.; Goodfellow, J.M. Molecular dynamics simulations of oligonucleotides in solution: Visualisation of intrinsic curvature. *J. Comput. Aided Mol. Des.* **1994**, *8*, 307–322.
50. Thompson, A.S.; Hurley, L.H. Solution conformation of a bizelesin A-tract duplex adduct: DNA-DNA cross-linking of an A-tract straightens out bent DNA. *J. Mol. Biol.* **1995**, *252*, 86–107.
51. Diekmann, S.; Mazzarelli, J.M.; McLaughlin, L.W.; von Kitzing, E.; Travers, A.A. DNA curvature does not require bifurcated hydrogen bonds or pyrimidine methyl groups. *J. Mol. Biol.* **1992**, *225*, 729–738.
52. Haran, T.E.; Kahn, J.D.; Crothers, D.M. Sequence elements responsible for DNA curvature. *J. Mol. Biol.* **1994**, *244*, 135–143.
53. Bruckner, I.; Sanchez, R.; Suck, D.; Pongor, S. Sequence-dependent bending propensity of DNA as revealed by DNase I-parameters for trinucleotides. *EMBO J.* **1995**, *14*, 1812–1818.
54. Bruckner, I.; Sanchez, R.; Suck, D.; Pongor, S. Trinucleotide models for DNA bending propensity—comparison of models based on DNase I digestion and nucleosome packaging data. *J. Biomol. Struct. Dyn.* **1995**, *13*, 309–317.
55. Drew, H.R.; Travers, A.A. DNA bending and its relation to nucleosome positioning. *J. Mol. Biol.* **1985**, *186*, 773–790.
56. Satchwell, S.C.; Drew, H.R.; Travers, A.A. Sequence periodicities in chicken nucleosome core DNA. *J. Mol. Biol.* **1986**, *191*, 659–675.
57. Wolffe, A.P.; Drew, H.R. DNA structure: implications for chromatin structure and function. In: *Chromatin Structure and Gene Expression* (Elgin, S.C.R., ed.). IRL Press, Oxford, 1995, pp. 27–48.
58. Hunter, C.A. Sequence-dependent DNA structure. The role of base stacking interactions. *J. Mol. Biol.* **1993**, *230*, 1025–1054.
59. Sprous, D.; Zacharias, W.; Wood, Z.A.; Harvey, S.C. Dehydrating agents sharply reduce curvature in DNAs containing A tracts. *Nucleic Acids Res.* **1995**, *23*, 1816–1821.
60. Dickerson, R.E.; Goodsell, D.; Kopka, M.L. MPD and DNA bending in crystals and in solution. *J. Mol. Biol.* **1996**, *256*, 108–125.
61. Calladine, C.R.; Drew, H.R. *Understanding DNA—The Molecule and How It Works*. Academic Press, San Diego, 1992.
62. Ner, S.S.; Travers, A.A.; Churchill, M.E.A. Harnessing the writhe—a role for DNA chaperones in nucleoprotein complex formation. *Trends Biochem. Sci.* **1994**, *19*, 185–187.
63. Travers, A.A.; Ner, S.S.; Churchill, M.E.A. DNA chaperones—a solution to a persistence problem. *Cell* **1994**, *77*, 167–169.
64. Bracco, L.; Kotlarz, D.; Kolb, A.; Diekmann, S.; Buc, H. Synthetic curved DNA sequences can act as transcriptional activators in *Escherichia coli*. *EMBO J.* **1989**, *8*, 4289–4296.
65. Yamada, H.; Muramatsu, S.; Mizuno, T. An *Escherichia coli* protein that preferentially binds to sharply curved DNA. *J. Biochem. (Tokyo)* **1990**, *108*, 420–425.
66. Tippner, D.; Afflerbach, H.; Bradaczek, C.; Wagner, R. Evidence for a regulatory function of the histone-like *Escherichia coli* protein H-NS in ribosomal-RNA synthesis. *Mol. Microbiol.* **1994**, *11*, 589–604.
67. Barcelo, F.; Muzard, G.; Mendoza, R.; Revet, B.; Roques, B.P.; Le Pecq, J.B. Removal of DNA curving by DNA ligands: gel electrophoresis study. *Biochemistry* **1991**, *30*, 4863–4873.
68. Owen-Hughes, T.A.; Pavitt, G.D.; Santos, D.S.; Sidebotham, J.M.; Hulton, C.; Hilton, J. et al. The chromatin-associated protein H-NS interacts with curved DNA to influence DNA topology and gene-expression. *Cell* **1992**, *71*, 255–265.
69. Dersch, P.; Schmidt, K.; Bremer, K. Synthesis of the *Escherichia coli* K-12 nucleoid-associated DNA-binding protein H-NS is subjected to growth-phase control and autoregulation. *Mol. Microbiol.* **1993**, *8*, 875–889.
70. Falconi, M.; Higgins, N.; Spurio, R.; Pon, C.; Gualerzi, C.O. Expression of the gene encoding the major bacterial nucleoid protein H-NS is subject to transcriptional auto-repression. *Mol. Microbiol.* **1993**, *10*, 273–282.

71. Ueguchi, C.; Kakeda, M.; Mizuno, T. Autoregulatory expression of the *Escherichia coli* hns gene encoding a nucleoid protein: H-NS functions as a repressor of its own transcription. *Mol. Gen. Genet.* **1993**, *236*, 171–178.
72. Zuber, F.; Kotlarz, D.; Rimsky, S.; Buc, H. Modulated expression of promoters containing upstream curved DNA sequences by the *Escherichia coli* nucleoid protein H-NS. *Mol. Microbiol.* **1994**, *12*, 231–240.
73. Yamada, H.; Yoshida, T.; Tanaka, K.; Sasakawa, C.; Mizuno, T. Molecular analysis of the *Escherichia coli* gene encoding a DNA-binding protein, which preferentially recognises curved DNA sequences. *Mol. Gen. Genet.* **1991**, *230*, 332–336.
74. Tanaka, K.; Muramatsu, S.; Yamada, H.; Mizuno, T. Systematic characterization of curved DNA segments randomly cloned from *Escherichia coli* and their functional significance. *Mol. Gen. Genet.* **1991**, *226*, 367–376.
75. Lucht, J.M.; Dersch, P.; Kempf, B.; Bremer, E. Interactions of the nucleoid-associated DNA-binding protein H-NS with the regulatory region of the osmotically controlled proU operon of *Escherichia coli*. *J. Biol. Chem.* **1994**, *269*, 6578–6586.
76. Mazin, A.; Timchenko, T.; Demurica, J.M.; Schreiber, V.; Angulo, J.F.; Demurcia, G. et al. Kin17, a mouse nuclear zinc-finger protein that binds preferentially to curved DNA. *Nucleic Acids Res.* **1994**, *22*, 4335–4341.
77. Gadaleta, G.; Delia, D.; Saccone, C.; Pepe, G. A 67-kDa protein-binding to the curved DNA in the main regulatory region of the rat mitochondrial genome. *Gene* **1995**, *160*, 229–234.
78. Lilley, D.M.J. HMG has DNA wrapped up. *Nature* **1992**, *357*, 282–283.
79. Schroth, G.P.; Siino, J.S.; Cooney, C.A.; Th'ng, J.P.; Ho, P.S.; Bradbury, E.M. Intrinsically bent DNA flanks both sides of an RNA polymerase I transcription start site. Both regions display novel electrophoretic mobility. *J. Biol. Chem.* **1992**, *267*, 9958–9964.
80. Schultz, S.C.; Shields, G.C.; Steitz, T.A. Crystal structure of a CAP-DNA complex: The DNA is bent by 90 degrees. *Science* **1991**, *253*, 1001–1007.
81. Ramstein, J.; Lavery, R. Energetic coupling between DNA bending and base pair opening. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 7231–7235.
82. Parvin, J.D.; McCormick, R.J.; Sharp, P.A.; Fisher, D.E. Pre-bending of a promoter sequence enhances affinity for the TATA-binding factor. *Nature* **1995**, *373*, 724–727.
83. Travers, A.A.; Klug, A. The bending of DNA in nucleosomes and its wider implications. *Philos. Trans. R. Soc. Lond. Biol.* **1987**, *317*, 537–561.
84. Collis, C.M.; Molloy, P.L.; Both, G.W.; Drew, H.R. Influence of the sequence-dependent flexure of DNA on transcription in *E. coli*. *Nucleic Acids Res.* **1989**, *17*, 9447–9468.
85. Figueroa, N.; Wills, N.; Bossi, L. Common sequence determinants of the response of a prokaryotic promoter to DNA bending and supercoiling. *EMBO J.* **1991**, *10*, 941–949.
86. Perez-Martin, J.; Rojo, F.; Delorenzo, V. Promoters responsive to DNA bending—a common theme in prokaryotic gene expression. *Microbiol. Rev.* **1994**, *58*, 268–290.
87. Lamond, A.I.; Travers, A.A. Requirement for an upstream element for optimal transcription of a bacterial tRNA gene. *Nature* **1983**, *305*, 248–250.
88. Deuschle, U.; Krammerer, W.; Gentz, R.; Bujard, H. Promoters of *Escherichia coli*: a hierarchy of *in vivo* strength indicates alternate structures. *EMBO J.* **1986**, *5*, 2987–2994.
89. Knaus, R.; Bujard, H. Principles governing the activity of *E. coli* promoters. In: *Nucleic Acids and Molecular Biology* (Eckstein, F.; Lilley, D.M.J.; eds.). Springer-Verlag, Berlin, 1990, pp. 110–112.
90. Plaskon, R.R.; Wartell, R.M. Sequence distributions associated with DNA curvature are found upstream of strong *E. coli* promoters. *Nucleic Acids Res.* **1987**, *15*, 785–796.
91. Ohyama, T.; Hirota, Y. A possible function of DNA curvature in transcription. *Nucleic Acids Symp. Ser.* **1993**, *29*, 153–154.
92. Kim, J.; Klooster, S.; Shapiro, D.J. Intrinsically bent DNA in a eukaryotic transcription factor recognition sequence potentiates transcription activation. *J. Biol. Chem.* **1995**, *270*, 1282–1288.

93. Goodman, S.D.; Nash, H.A. Functional replacement of a protein-induced bend in a DNA recombination site. *Nature* **1989**, *341*, 251–254.
94. Tanaka, K.-I.; Ueguchi, C.; Mizuno, T. Importance of stereospecific positioning of the upstream cis-acting DNA element containing curved DNA structure for the functioning of the *Escherichia coli* proV promoter. *Biosci. Biotechnol. Biochem.* **1994**, *58*, 1097–1101.
95. Ross, W.; Gosink, K.; Salomon, J.; Igarashi, K.; Zhou, C.; Ishihama, A. et al. A third recognition element in bacterial promoters: DNA binding by the alpha subunit of RNA polymerase. *Science* **1993**, *262*, 1407–1413.
96. Rao, L.; Ross, W.; Appleman, J.A.; Gaal, T.; Leirmo, S.; Schlax, P.J., et al. Factor independent activation of rrnB1 P1—an “extended promoter” with an upstream element that dramatically increases promoter strength. *J. Mol. Biol.* **1994**, *235*, 1421–1435.
97. Jordi, B.J.A.M.; Owen-Hughes, T.A.; Hulton, C.S.J.; Higgins, C.F. DNA twist, flexibility and transcription of the osmoregulated proU promoter of *Salmonella typhimurium*. *EMBO J.* **1995**, *22*, 5690–5700.
98. Nadal, M.; Mirambeau, G.; Forterre, P.; Reiter, W.-D.; Duguet, M. Positively supercoiled DNA in a virus-like particle of an archaebacterium. *Nature* **1986**, *321*, 256–258.
99. Spengler, S.J.; Stasiak, A.; Cozzarelli, N.R. The stereostructure of knots and catenanes produced by phage lambda integrative recombination: Implications for mechanism and DNA structure. *Cell* **1985**, *42*, 325–334.
100. Wasserman, S.A.; Cozzarelli, N.R. Biochemical topology: applications to DNA recombination and replication. *Science* **1986**, *232*, 951–960.
101. Adams, D.E.; Shekhtman, E.M.; Zechiedrich, E.L.; Schmid, M.B.; Cozzarelli, N.R. The role of topoisomerase IV in partitioning bacterial replicons and the structure of catenated intermediates in DNA replication. *Cell* **1992**, *71*, 277–288.
102. Depew, R.E.; Wang, J.C. Conformational fluctuations of the DNA helix. *Proc. Natl. Acad. Sci. USA* **1975**, *72*, 4275–4279.
103. Pulleyblank, D.E.; Shure, M.; Tang, D.; Vinograd, J.; Vosberg, H. Action of nick-closing enzyme on supercoiled and non-supercoiled closed circular DNA: formation of a Boltzmann distribution of topological isomers. *Proc. Natl. Acad. Sci. USA* **1975**, *72*, 4280–4284.
104. Drlica, K. Biology of bacterial deoxyribonucleic acid topoisomerases. *Microbiol. Rev.* **1984**, *48*, 273–289.
105. Drlica, K. Bacterial topoisomerases and the control of DNA supercoiling. *Trends Genet.* **1990**, *6*, 433–437.
106. Gellert, M. DNA topoisomerases. *Annu. Rev. Biochem.* **1981**, *50*, 879–910.
107. Wang, J.C.; Lynch, A.S. Transcription and DNA supercoiling. *Curr. Opin. Genet. Dev.* **1993**, *3*, 764–768.
108. Chung, J.H.; Whiteley, M.; Felsenfeld, G. A 5' element of the chicken β -globin domain serves as an insulator in human erythroid cells and protects against position effect in *Drosophila*. *Cell* **1993**, *74*, 505–514.
109. Drlica, K. Control of bacterial DNA supercoiling. *Mol. Microbiol.* **1992**, *6*, 425–433.
110. Hsieh, L.S.; Burger, R.M.; Drlica, K. Bacterial DNA supercoiling and [ATP]/[ADP] changes associated with a transition to anaerobic growth. *J. Mol. Biol.* **1991**, *219*, 1–8.
111. Pruss, G.J.; Drlica, K. Topoisomerase I mutants: The gene on pBR322 that encodes resistance to tetracycline affects plasmid DNA supercoiling. *Proc. Natl. Acad. Sci. USA* **1986**, *83*, 8952–8956.
112. Lockshon, D.; Morris, D.R. Positively supercoiled plasmid DNA is produced by treatment of *Escherichia coli* with DNA gyrase inhibitors. *Nucleic Acids Res.* **1983**, *11*, 2999–3018.
113. Liu, L.F.; Wang, J.C. Supercoiling of the DNA template during transcription. *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 7024–7027.
114. Esposito, F.; Sinden, R.R. DNA supercoiling and eukaryotic gene expression. *Ox. Surv. Eukaryot. Genes* **1988**, *5*, 1–50.

115. Freeman, L.A.; Garrard, W.T. DNA supercoiling in chromatin structure and gene expression. *Crit. Rev. Eukaryot Gene Expr.* **1992**, *2*, 165–209.
116. Horwitz, M.S.Z.; Loeb, L.A. An *E. coli* promoter that regulates transcription by DNA superhelix-induced cruciform extrusion. *Science* **1988**, *241*, 703–705.
117. McClure, W.P. Mechanism and control of transcription in prokaryotes. *Annu. Rev. Biochem.* **1985**, *54*, 171–204.
118. Pearson, C.E.; Zorbas, H.; Price, G.B.; Zannis-Hadjopoulos, M. Inverted repeats, stem-loops, and cruciforms: significance for initiation of DNA replication. *J. Cell. Biochem.* **1996**, *63*, 1–22.
119. Murchie, A.I.; Lilley, D.M. The mechanism of cruciform formation in supercoiled DNA: initial opening of central basepairs in salt-dependent extrusion. *Nucleic Acids Res.* **1987**, *15*, 9641–9654.
120. Zheng, G.; Sinden, R.R. Effect of base composition at the center of inverted repeated DNA sequences on cruciform transitions in DNA. *J. Biol. Chem.* **1988**, *263*, 5356–6461.
121. Courey, A.J.; Wang, J.C. Influence of DNA sequence and supercoiling on the process of cruciform formation. *J. Mol. Biol.* **1988**, *202*, 35–43.
122. Lilley, D.M. The inverted repeat as a recognizable structural feature in supercoiled DNA. *Proc. Natl. Acad. Sci. USA* **1980**, *77*, 6468–6472.
123. Panayotatos, N.; Wells, R.D. Cruciform structures in supercoiled DNA. *Nature* **1981**, *289*, 466–470.
124. Mizuuchi, K.; Mizuuchi, M.; Gellert, M. Cruciform structures in palindromic DNA are favored by DNA supercoiling. *J. Mol. Biol.* **1982**, *156*, 229–243.
125. Gellert, M.; O'Dea, M.H.; Mizuuchi, K. Slow cruciform transitions in palindromic DNA. *Proc. Natl. Acad. Sci. USA* **1983**, *80*, 5545–5549.
126. Singleton, C.K. Effects of salts, temperature, and stem length on supercoil-induced formation of cruciforms. *J. Biol. Chem.* **1983**, *258*, 7661–7668.
127. Sinden, R.R.; Pettijohn, D.E. Cruciform transitions in DNA. *J. Biol. Chem.* **1984**, *259*, 6593–6600.
128. Sullivan, K.M.; Lilley, D.M. Influence of cation size and charge on the extrusion of a salt-dependent cruciform. *J. Mol. Biol.* **1987**, *193*, 397–404.
129. Lilley, D.M. The kinetic properties of cruciform extrusion are determined by DNA. *Nucleic Acids Res.* **1985**, *13*, 1443–1465.
130. Sullivan, K.M.; Lilley, D.M. A dominant influence of flanking sequences on a local structural transition. *Cell* **1986**, *47*, 817–827.
131. Sullivan, K.M.; Murchie, A.I.; Lilley, D.M. Long range structural communication between sequences in supercoiled DNA. Sequence dependence of contextual influence on cruciform extrusion mechanism. *J. Biol. Chem.* **1988**, *263*, 13074–13082.
132. Bowater, R.; Aboul-ela, F.; Lilley, D.M.J. Large-scale stable opening of supercoiled DNA in response to temperature and supercoiling in (A + T)-rich regions that promote low-salt cruciform extrusion. *Biochemistry* **1991**, *30*, 11495–11506.
133. Morales, N.M.; Cobourn, S.D.; Muller, U.R. Effect of *in vitro* transcription on cruciform stability. *Nucleic Acids Res.* **1990**, *18*, 2777–2782.
134. Zheng, G.; Kochel, T.; Hoepfner, R.W.; Timmons, S.E.; Sinden, R.R. Torsionally tuned cruciform and Z-DNA probes for measuring unrestrained supercoiling at specific sites in DNA of living cells. Appendix: estimation of superhelical density *in vivo* from analysis of the level of cruciforms existing in living cells. *J. Mol. Biol.* **1991**, *221*, 107–129.
135. Lilley, D.M.J.; Clegg, R.M. The structure of the four-way junction in DNA. *Annu. Rev. Biophys. Biomol. Struct.* **1993**, *22*, 299–328.
136. Wemmer, D.E.; Wand, A.J.; Seeman, N.C.; Kallenbach, N.R. NMR analysis of DNA junctions: imino proton NMR studies of individual arms and intact junction. *Biochemistry* **1985**, *24*, 5745–5749.

137. Marky, L.A.; Kallenbach, N.R.; McDonough, K.A.; Seeman, N.C.; Breslauer, K.J. Melting behaviour of a DNA junction structure: a calorimetric and spectroscopic study. *Biopolymers* **1987**, *26*, 1621–1634.
138. Murchie, A.I.; Clegg, R.M.; von Kitzing, E.; Duckett, D.R.; Diekmann, S.; Lilley, D.M. Fluorescence energy transfer shows that the four-way DNA junction is a right-handed cross of anti-parallel molecules. *Nature* **1989**, *341*, 763–766.
139. Churchill, M.E.A.; Tullius, T.D.; Kallenbach, N.R.; Seeman, N.C. A Holliday recombination intermediate is twofold symmetric. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 4653–4656.
140. Kimball, A.; Guo, Q.; Lu, M.; Cunningham, R.P.; Kallenbach, N.R.; Seeman, N.C., et al. Construction and analysis of parallel and antiparallel Holliday junctions. *J. Biol. Chem.* **1990**, *265*, 6544–6547.
141. Murchie, A.I.H.; Portugal, J.; Lilley, D.M.J. Cleavage of a four-way DNA junction by a restriction enzyme spanning the point of strand exchange. *EMBO J.* **1991**, *10*, 713–718.
142. Chen, J.H.; Churchill, M.E.A.; Tullius, T.D.; Kallenbach, N.R.; Seeman, N.C. Construction and analysis of monomobile DNA junctions. *Biochemistry* **1988**, *27*, 6032–6038.
143. Duckett, D.R.; Lilley, D.M.J. Effects of base mismatches on the structure of the four-way DNA junction. *J. Mol. Biol.* **1991**, *221*, 147–161.
144. Duckett, D.R.; Murchie, A.I.; Lilley, D.M. The role of metal ions in the conformation of the four-way DNA junction. *EMBO J.* **1990**, *9*, 583–590.
145. Duckett, D.R.; Murchie, A.I.H.; Diekmann, S.; von Kitzing, E.; Kemper, B.; Lilley, D.M.J. The structure of the Holliday junction, and its resolution. *Cell* **1988**, *55*, 79–89.
146. Clegg, R.M.; Murchie, A.I.H.; Zechel, A.; Carlberg, C.; Diekmann, S.; Lilley, D.M.J. Fluorescence resonance energy transfer analysis of the structure of the four-way DNA junction. *Biochemistry* **1992**, *31*, 4846–4856.
147. Duckett, D.R.; Murchie, A.I.H.; Clegg, R.M.; Zechel, A.; von Kitzing, E.; Diekmann, S., et al. The structure of the Holliday junction. In: *Structure and Methods* (Sarma, M.H.; Sarma, R.H.; Eds.), Adenine Press, Albany, 1990, pp. 157–181.
148. Clegg, R.M.; Murchie, A.I.H.; Lilley, D.M.J. The solution structure of the four-way DNA junction at low salt conditions: A fluorescence resonance energy transfer analysis. *Biophys. J.* **1994**, *66*, 99–109.
149. Eis, P.S.; Millar, D.P. Conformational distributions of a four-way DNA junction revealed by time-fluorescence energy transfer. *Biochemistry* **1993**, *32*, 13852–13860.
150. Gough, G.W.; Lilley, D.M. DNA bending induced by cruciform formation. *Nature* **1985**, *313*, 154–156.
151. Pil, P.M.; Lippard, S.J. Specific binding of chromosomal protein HMG1 to DNA damaged by the anticancer drug cisplatin. *Science* **1992**, *256*, 234–237.
152. Scholten, P.M.; Nordheim, A. Diethyl pyrocarbonate: a chemical probe for DNA cruciforms. *Nucleic Acids Res.* **1986**, *14*, 3981–3993.
153. Furlong, J.C.; Lilley, D.M. Highly selective chemical modification of cruciform loops by diethyl pyrocarbonate. *Nucleic Acids Res.* **1986**, *14*, 3995–4007.
154. Blommers, M.J.; Van de Ven, F.J.M.; Van der Marel, G.A.; Van Boom, J.H.; Hilbers, C.W. The three-dimensional structure of a DNA hairpin in solution two dimensional NMR studies and structural analysis of d(ATCTTATTATAGGAT). *Eur. J. Biochem.* **1991**, *201*, 33–51.
155. Haasnoot, C.A.G.; Hilbers, C.W.; Van der Marel, G.A.; Van Boom, J.H.; Singh, U.C.; Pattabiraman, N., et al. On loop-folding in nucleic acid hairpin-type structures. *J. Biomol. Struct. Dyn.* **1986**, *3*, 843–857.
156. Hirao, I.; Kawai, G.; Yoshizawa, S.; Nishimura, Y.; Ishido, Y.; Watanabe, K., et al. Most compact hairpin-turn structure exerted by a short DNA fragment, d(GCGAAGC) in solution: an extraordinarily stable structure resistant to nucleases and heat. *Nucleic Acids Res.* **1994**, *22*, 576–285.
157. Zhou, N.; Vogel, H.J. Two dimensional NMR and restrained molecular dynamics studies of the hairpin d(T₈C₄A₈): Detection of an extraloop cytosine. *Biochemistry* **1993**, *32*, 637–645.

158. Williamson, J.R.; Boxer, S.G. Multinuclear NMR studies of DNA hairpins. 2. Sequence-dependent structural variations. *Biochemistry* **1989**, *28*, 2831–2836.
159. Germann, M.W.; Kalisch, B.W.; Lundberg, P.; Vogel, H.J.; van de Sande, J. Perturbation of DNA hairpins containing the EcoRI recognition site by hairpin loops of varying size and composition: physical (NMR and UV) and enzymatic (EcoRI) studies. *Nucleic Acids Res.* **1990**, *18*, 1489–1498.
160. Pearson, C.E.; Zannis-Hadjopoulos, M.; Price, G.B.; Zorbas, H. A novel type of interaction between cruciform DNA and a cruciform binding protein from HeLa cells. *EMBO J.* **1995**, *14*, 1571–1580.
161. Holliday, R. A mechanism for gene conversion in fungi. *Genet. Res.* **1964**, *5*, 282–304.
162. Timsit, Y.; Moras, D. Groove-backbone interaction in B-DNA implication for DNA condensation and recombination. *J. Mol. Biol.* **1991**, *221*, 919–940.
163. Panyutin, I.G.; Hsieh, P. Formation of a single base mismatch impedes spontaneous DNA branch migration. *J. Mol. Biol.* **1993**, *230*, 413–424.
164. Panyutin, I.G.; Hsieh, P. The kinetics of spontaneous DNA branch migration. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 2021–2025.
165. Panyutin, I.G.; Biswas, I.; Hsieh, P. A pivotal role for the structure of the Holliday junction in DNA branch migration. *EMBO J.* **1995**, *14*, 1819–1826.
166. Mulrooney, S.B.; Fishel, R.A.; Hejna, J.A.; Warner, R.C. Preparation of figure 8 and cruciform DNAs and their use in studies of the kinetics of branch migration. *J. Biol. Chem.* **1996**, *271*, 9648–9659.
167. Evans, D.H.; Kolodner, R. Effect of DNA structure and nucleotide sequence on Holliday junction resolution by a *Saccharomyces cerevisiae* endonuclease. *J. Mol. Biol.* **1987**, *262*, 969–980.
168. Evans, D.H.; Kolodner, R. Construction of a synthetic Holliday junction analog and characterization of its interaction with a *Saccharomyces cerevisiae* endonuclease that cleaves Holliday junctions. *J. Biol. Chem.* **1987**, *262*, 9160–9165.
169. Passanti, C.; Davies, B.; Ford, M.; Fried, M. Structure of an inverted duplication formed as a first step in a gene amplification event: implications for a model of gene amplification. *EMBO J.* **1987**, *6*, 1697–1703.
170. Mueller, J.E.; Newton, C.J. Resolution of Holliday junction analogs by T4 endonuclease VII can be directed by substrate structure. *J. Biol. Chem.* **1990**, *265*, 13918–13924.
171. Dunderdale, H.J.; Benson, F.E.; Parsons, C.A.; Sharples, G.J.; Lloyd, R.G.; West, S.C. Formation and resolution of recombination intermediates by *E. coli* RecA and RuvC proteins. *Nature* **1991**, *354*, 5506–5510.
172. Bennett, R.J.; Dunderdale, H.J.; West, S.C. Resolution of Holliday junctions by RuvC resolvase: cleavage specificity and DNA distortion. *Cell* **1993**, *74*, 1021–1031.
173. Waga, S.; Mizuno, S.; Yoshida, M. Chromosomal protein HMG1 removes the transcriptional block caused by the cruciform in supercoiled DNA. *J. Biol. Chem.* **1990**, *265*, 19424–19428.
174. Shiba, T.; Iwasaki, H.; Nakata, A.; Shinagawa, H. SOS-inducible DNA repair proteins, RuvA and RuvB, of *Escherichia coli*: functional interactions between RuvA and RuvB for ATP hydrolysis and renaturation of the cruciform structure in supercoiled DNA. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 8445–8449.
175. Adams, D.E.; West, S.C. Relaxing and unwinding on Holliday: DNA helicase mediated branch migration. *Mutat. Res.* **1995**, *337*, 149–159.
176. Lilley, D.M. *In vivo* consequences of plasmid topology. *Nature* **1981**, *292*, 380–382.
177. Palecek, E. Local supercoil-stabilized DNA structures. *CRC Crit. Rev. Biochem. Mol. Biol.* **1991**, *26*, 151–226.
178. Furlong, J.C.; Sullivan, K.M.; Murchie, A.I.; Gough, G.W.; Lilley, D.M. Localized chemical hyperreactivity in supercoiled DNA: evidence for base unpairing in sequences that induce low-salt cruciform extrusion. *Biochemistry* **1989**, *28*, 2009–2017.

179. Lilley, D.M.; Kemper, B. Cruciform-resolvase interactions in supercoiled DNA. *Cell* **1984**, *36*, 413–422.
180. Sinden, R.R.; Broyles, S.S.; Pettijohn, D.E. Perfect palindromic lac operator DNA sequence exists as a stable cruciform structure in supercoiled DNA *in vitro* but not *in vivo*. *Proc. Natl. Acad. Sci. USA* **1983**, *80*, 1797–1801.
181. Frappier, L.; Price, G.B.; Martin, R.G.; Zannis-Hadjopoulos, M. Monoclonal antibodies to cruciform DNA structures. *J. Mol. Biol.* **1987**, *193*, 751–758.
182. Cimino, G.D.; Gamper, H.B.; Isaacs, S.T.; Hearst, J.E. Psoralens as photoactive probes of nucleic acid structure and function: organic chemistry, photochemistry, and biochemistry. *Annu. Rev. Biochem.* **1985**, *54*, 1151–1193.
183. Ussery, D.W.; Hoepfner, R.W.; Sinden, R.R. Probing DNA structure with psoralen *in vitro*. *Methods Enzymol.* **1992**, *212*, 242–262.
184. Sinden, R.R.; Ussery, D.W. Analysis of DNA structure *in vivo* using psoralen photobinding: measurement of supercoiling, topological domains, and DNA-protein interactions. *Methods Enzymol.* **1992**, *212*, 319–335.
185. Lis, J.T.; Neckameyer, W.; Mirault, M.-E.; Artavanis-Tsakonas, S.; Lall, P.; Martin, P.; et al. DNA sequences flanking the starts of the hsp70 and alpha beta heat shock genes are homologous. *Dev. Biol.* **1981**, *83*, 291–300.
186. Panayotatos, N.; Fontaine, A. A native cruciform DNA structure probed in bacteria by recombinant T7 endonuclease. *J. Biol. Chem.* **1987**, *262*, 11364–11368.
187. McClellan, J.A.; Boublíkova, P.; Palecek, E.; Lilley, D.M. Superhelical torsion in cellular DNA responds directly to environmental and genetic factors. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 8373–8377.
188. Haniford, D.B.; Pulleyblank, D.E. Transition of a cloned d(AT)_n-d(AT)_n tract to a cruciform *in vivo*. *Nucleic Acids Res.* **1985**, *13*, 4343–4361.
189. Ward, G.K.; McKenzie, R.; Zannis-Hadjopoulos, M.; Price, G.B. The dynamic distribution and quantification of DNA cruciforms in eukaryotic nuclei. *Exp. Cell. Res.* **1990**, *188*, 235–246.
190. Ward, G.K.; Shihab-el-Deen, A.; Zannis-Hadjopoulos, M.; Price, G.B. DNA cruciforms and the nuclear supporting structure. *Exp. Cell. Res.* **1991**, *195*, 92–98.
191. Jackson, D.A.; Dickinson, P.; Cook, P.R. The size of chromatin loops in HeLa cells. *EMBO J.* **1990**, *9*, 567–571.
192. Burkholder, G.D.; Latimer, L.J.P.; Lee, J.S. Immunofluorescent staining of mammalian nuclei and chromosomes with a monoclonal antibody to triplex DNA. *Chromosoma* **1988**, *97*, 185–192.
193. Wittig, B.; Dorbic, T.; Rich, A. Transcription is associated with Z-DNA formation in metabolically active permeabilized mammalian cell nuclei. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 2259–2263.
194. Wittig, B.; Dorbic, T.; Rich, A. The level of Z-DNA in metabolically active, permeabilized mammalian cell nuclei is regulated by torsional strain. *J. Cell. Biol.* **1989**, *108*, 755–764.
195. Klein, W.J.; Beiser, S.M.; Erlanger, B.F. Nuclear fluorescence employing antinucleoside immunoglobins. *J. Exp. Med.* **1967**, *125*, 61–70.
196. Tan, E.M.; Lerner, R.A. An immunological study of the fates of nuclear and nucleolar macromolecules during the cell cycle. *J. Mol. Biol.* **1972**, *68*, 107–114.
197. Collins, J.M. Deoxyribonucleic acid structure in human diploid fibroblasts stimulated to proliferate. *J. Biol. Chem.* **1977**, *252*, 141–147.
198. Conrad, M.N.; Newlon, C.S. Stably denatured regions of chromosomal DNA from the cdc 2 *Saccharomyces cerevisiae* cell cycle mutant. *Mol. Cell. Biol.* **1983**, *3*, 1665–1669.
199. Stillman, B. Smart machines at the replication fork. *Cell* **1994**, *78*, 725–728.
200. DePamphilis ML. Origins of DNA replication in metazoan chromosomes. *J. Biol. Chem.* **1993**, *268*, 1–4.
201. Hagerman, P.J. Sequence-directed curvature of DNA. *Annu. Rev. Biochem.* **1990**, *59*, 755–781.

202. Umek, R.M.; Kowalski, D. The ease of DNA unwinding as a determinant of initiation at yeast replication origins. *Cell* **1988**, *52*, 559–567.
203. Natale, D.A.; Umek, R.M.; Kowalski, D. Ease of unwinding is a conserved property of yeast replication origins. *Nucleic Acids Res.* **1993**, *21*, 555–560.
204. Kowalski, D.; Eddy, M.J. The DNA unwinding element: a novel, cis-acting component that facilitates opening of the *Escherichia coli* replication origin. *EMBO J.* **1989**, *8*, 4335–4344.
205. Umek, R.M.; Kowalski, D. Yeast regulatory sequences preferentially adopt a non-B conformation in supercoiled DNA. *Nucleic Acids Res.* **1987**, *15*, 4467–4480.
206. Lin, S.; Kowalski, D. DNA helical instability facilitates initiation at the SV40 replication origin. *J. Mol. Biol.* **1994**, *235*, 496–507.
207. Hay, R.T.; DePamphilis, M.L. Initiation of SV40 DNA replication in vivo: location and structure of 5' ends of DNA synthesized in the ori region. *Cell* **1982**, *28*, 767–779.
208. Hay, R.T.; Hendrickson, E.A.; DePamphilis, M.L. Sequence specificity for the initiation of RNA-primed simian virus 40 DNA synthesis in vivo. *J. Mol. Biol.* **1984**, *175*, 131–157.
209. Yagil, G. Parasitic structures of DNA and their role in DNA unwinding. *CRC Crit. Rev. Biochem. Mol. Biol.* **1991**, *26*, 475–559.
210. Murchie, A.I.H.; Bowater, R.; Aboul-ela, F.; Lilley, D.M.J. Helix opening transitions in supercoiled DNA. *Biochim. Biophys. Acta* **1992**, *1131*, 1–15.
211. Williams, L.; Kowalski, D. Easily unwound DNA sequences and hairpin structures in the Epstein-Barr virus origin of plasmid replication. *J. Virol.* **1993**, *67*, 2707–2715.
212. Weller, S.K.; Spadaro, A.; Schaffer, J.E.; Murray, A.W.; Maxam, A.M.; Schaffer, P.A. Cloning, sequencing, and functional analysis of oriL, a herpes simplex virus type 1 origin of DNA synthesis. *Mol. Cell Biol.* **1985**, *5*, 930–942.
213. Lockshon, D.; Galloway, D.A. Cloning and characterization of oriL2, a large palindromic DNA replication origin of herpes simplex virus type 2. *J. Virol.* **1986**, *58*, 513–521.
214. Noirot, P.; Bargoniotti, J.; Novick, R.P. Initiation of rolling-circle replication in pT181 plasmid: initiator protein enhances cruciform extrusion at the origin. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 8560–8564.
215. Frappier, L.; Price, G.B.; Martin, R.G.; Zannis-Hadjopoulos, M. Characterization of the binding specificity of two anticruciform DNA monoclonal antibodies. *J. Biol. Chem.* **1989**, *264*, 334–341.
216. Steinmetzer, K.; Zannis-Hadjopoulos, M.; Price, G.B. Anti-cruciform monoclonal antibody and cruciform DNA interaction. *J. Mol. Biol.* **1995**, *254*, 29–37.
217. Zannis-Hadjopoulos, M.; Frappier, L.; Khoury, M.; Price, G.B. Effect of anti-cruciform DNA monoclonal antibodies on DNA replication. *EMBO J.* **1988**, *7*, 1837–1844.
218. Fried, M.; Feo, S.; Heard, E. The role of inverted duplication in the generation of gene amplification in mammalian cells. *Biochim. Biophys. Acta* **1991**, *1090*, 143–155.
219. Windle, B.E.; Wahl, G.M. Molecular dissection of mammalian gene amplification: new mechanistic insights revealed by analysis of very early events. *Mutat. Res.* **1992**, *276*, 199–224.
220. Legouy, E.; Fossar, N.; Lhomond, G.; Brison, O. Structure of four amplified DNA novel joints. *Somat. Cell Mol. Genet.* **1989**, *15*, 309–320.
221. Hyrien, O. Large inverted duplications in amplified DNA of mammalian cells form hairpins *in vitro* upon DNA extraction but not *in vivo*. *Nucleic Acids Res.* **1989**, *17*, 9557–9569.
222. Esposito, F.; Sinden, R.R. Supercoiling in eukaryotic and prokaryotic DNA: changes in response to topological perturbation of SV40 in CV-1 cells and of plasmids in *E. coli*. *Nucleic Acids Res.* **1987**, *15*, 5105–5124.
223. Cohen, S.; Hassin, D.; Karby, S.; Lavi, S. Hairpin structures are the primary amplification products: a novel mechanism for generation of inverted repeats during gene amplification. *Mol. Cell Biol.* **1994**, *14*, 7782–7791.
224. Berko-Flint, Y.; Karby, S.; Hassin, D.; Lavi, S. Carcinogen-induced DNA amplification *in vitro*, overreplication of simian virus 40 origin region in extracts from carcinogen-treated CO60 cells. *Mol. Cell Biol.* **1990**, *10*, 75–83.

225. Aladjem, M.I.; Lavi, S. The mechanism of carcinogen-induced DNA amplification *in vivo* and *in vitro*. *Mutat. Res.* **1992**, *276*, 339–344.
226. Workman, J.L.; Taylor, C.A.; Kingston, R.E. Activation domains of stably bound GAL4 derivatives alleviate repression of promoters by nucleosomes. *Cell* **1991**, *64*, 533–544.
227. Cheng, L.; Kelly, T.J. The transcriptional activator nuclear factor 1 stimulates the replication of SV40 minichromosomes *in vivo* and *in vitro*. *Cell* **1989**, *59*, 541–551.
228. Simpson, R.T. Nucleosome positioning: Occurrence, mechanisms, and functional consequences. *Prog. Nucleic Acids Res. Mol. Biol.* **1990**, *40*, 143–184.
229. Weintraub, H. A dominant role for DNA secondary structure in forming hypersensitive structures in chromatin. *Cell* **1983**, *32*, 1191–1203.
230. Weintraub, H.; Cheng, P.F.; Conrad, K. Expression of transfected DNA depends on DNA topology. *Cell* **1986**, *46*, 115–122.
231. Nickol, J.; Martin, R.G. DNA stem-loop structures bind poorly to histone octamer cores. *Proc. Natl. Acad. Sci. USA* **1983**, *80*, 4669–4673.
232. Nobile, C.; Nickol, J.; Martin, R.G. Nucleosome phasing on a DNA fragment from the replication origin of simian virus 40 and rephasing upon cruciform formation of the DNA. *Mol. Cell Biol.* **1986**, *6*, 2916–2922.
233. Battistoni, A.; Leoni, L.; Sampaolesi, B.; Savino, M. Kinetic persistence of cruciform structures in reconstituted minichromosomes. *Biochim. Biophys. Acta* **1988**, *950*, 161–171.
234. Kotani, H.; Kmiec, E.B. DNA cruciforms facilitate *in vitro* strand transfer on nucleosomal templates. *Mol. Gen. Genet.* **1994**, *243*, 681–690.
235. Nickol, J.; Behe, M.; Felsenfeld, G. Effect of the B-Z transition in poly(dG-m⁵dC)poly(dG-m⁵dC) on nucleosome formation. *Biochemistry* **1982**, *79*, 1771–1775.
236. Miller, F.D.; Dixon, G.H.; Rattner, J.B.; van de Sande, J.H. Assembly and characterization of nucleosomal cores on B- vs. Z-form DNA. *Biochemistry* **1985**, *24*, 102–109.
237. Garner, M.M.; Felsenfeld, G. Effect of Z-DNA on nucleosome placement. *J. Mol. Biol.* **1987**, *196*, 581–590.
238. Ausio, J.; Zhou, G.; van Holde, K. A re-examination of the reported B-Z DNA transition in nucleosomes reconstituted with poly(dG-m⁵dC)poly(dG-m⁵dC). *Biochemistry* **1987**, *26*, 5595–5599.
239. Gross, D.S.; Huang, S.Y.; Garrard, W.T. Chromatin structure of the potential Z-forming sequence (dT-dG)_n × (dC-dA)_n. Evidence for an “alternating-B” conformation. *J. Mol. Biol.* **1985**, *183*, 251–265.
240. Lee-Chen, G.; Woodworth-Gutai, M. Simian virus 40 DNA replication: functional organization of regulatory elements. *Mol. Cell Biol.* **1986**, *6*, 3086–3093.
241. Chandrasekharappa, S.C.; Subramanian, K.N. Effects of position and orientation of the 72-base-pair-repeat transcriptional enhancer on replication from the simian virus 40 core origin. *J. Virol.* **1987**, *61*, 2973–2980.
242. Jakobovits, E.B.; Bratosin, S.; Aloni, Y. A nucleosome-free region in SV40 minichromosomes. *Nature* **1980**, *285*, 263–265.
243. Saragosti, S.; Moyne, G.; Yanif, M. Absence of nucleosomes in a fraction of SV40 chromatin between the origin of replication and the region coding for the late reader RNA. *Cell* **1980**, *20*, 65–73.
244. Solomon, M.J.; Varshavsky, A. A nuclease-hypersensitive region from *de novo* after chromosome replication. *Mol. Cell Biol.* **1987**, *7*, 3822–3825.
245. Nordheim, A.; Rich, A. Negatively supercoiled simian virus 40 DNA contains Z-DNA segments within transcriptional enhancer sequences. *Nature* **1983**, *303*, 674–679.
246. Azorin, F.; Rich, A. Isolation of Z-DNA binding proteins from SV40 minichromosomes: evidence for binding to the viral control region. *Cell* **1985**, *41*, 365–374.
247. Sage, E.; Leng, M. Conformational changes of poly(dG-dC)-poly(dG-dC) modified by the carcinogen *N*-acetoxy-*N*-acetyl-2-aminofluorene. *Nucleic Acids Res.* **1981**, *9*, 1241

248. Nordheim, A.; Herrera, R.E.; Rich, A. Binding of anti-Z-DNA antibodies to negatively supercoiled SV40 DNA. *Nucleic Acids Res.* **1987**, *15*, 1661–1677.
249. Muller, R.C.; Raphael, A.L.; Barton, J.K. Evidence for altered DNA conformations in the simian virus 40 genome: site-specific DNA cleavage by the chiral complex Λ-tris (4,7-dihenyl-1,10-phenanthroline) cobalt (III). *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 1764–1768.
250. Sheflin, L.G.; Kowalski, D. Altered DNA conformations detected by mung bean nuclease occur in promoter and terminator regions of supercoiled pBR322 DNA. *Nucleic Acids Res.* **1985**, *13*, 6137–6155.
251. Iacono-Conners, L.; Kowalski, D. Altered conformations in gene regulatory regions of torsionally stressed SV40 DNA. *Nucleic Acids Res.* **1986**, *14*, 8949–8962.
252. Muller, U.R.; Fitch, W.M. Evolutionary selection for perfect hairpin structures in viral DNAs. *Nature* **1982**, *298*, 582–585.
253. Greenberg, M.E.; Siegfried, Z.; Ziff, E.B. Mutation of the *c-fos* dyad symmetry element inhibits serum inducibility *in vivo* and nuclear regulatory factor binding *in vitro*. *Mol. Cell Biol.* **1987**, *7*, 1217–1225.
254. Martinez-Arias, A.; Yost, H.J.; Casadaban, M.J. Role of an upstream regulatory element in leucine repression of the *Saccharomyces cerevisiae* leu2 gene. *Nature* **1984**, *307*, 740–742.
255. Shuster, J.; Yu, J.; Cox, D.; Chan, R.V.L.; Smith, M.; Young, E. ADR1-mediated regulation of ADH2 requires an inverted repeat sequence. *Mol. Cell Biol.* **1986**, *6*, 1894–1902.
256. Schon, E.; Evans, T.; Welsh, J.; Efstratiadis, A. Conformation of promoter DNA: fine mapping of S1-hypersensitive sites. *Cell* **1983**, *35*, 837–848.
257. Glucksmann, M.A.; Markiewicz, P.; Malone, C.; Rothman-Denes, L.B. Specific sequences and a hairpin structure in the template strand are required for N4 virion RNA polymerase promoter recognition. *Cell* **1992**, *70*, 491–500.
258. Horwitz, M.S.Z.; Loeb, L.A. Structure function relationships in *Escherichia coli* promoter DNA. *Prog. Nucleic Acids Res. Mol. Biol.* **1990**, *38*, 137–164.
259. Markiewicz, P.; Malone, C.; Chase, J.W.; Rothman-Denes, L.B. *Escherichia coli* single-stranded DNA-binding protein is a supercoiled template-dependent transcriptional activator of N4 virion RNA polymerase. *Genes Dev.* **1992**, *6*, 2010–2019.
260. Wu, H.Y.; Shyy, S.H.; Wang, J.C.; Liu, L.F. Transcription generates positively and negatively supercoiled domains in the template. *Cell* **1988**, *53*, 433–440.
261. Wang, J.C.; Liu, L.F. DNA replication: topological aspects and the roles of DNA topoisomerases. In: *DNA Topology and Its Biological Effects* (Cozzarelli, N.R.; Wang, J.C.; Eds.). Cold Spring Harbor, NY, Cold Spring Harbor Laboratory, 1990, pp. 321–340.
262. Dayn, A.; Makhosyan, S.; Mirkin, S. Transcriptionally driven cruciform formation *in vivo*. *Nucleic Acids Res.* **1992**, *20*, 5991–5997.
263. Wang, J.C. DNA topoisomerases. *Annu. Rev. Biochem.* **1985**, *54*, 665–697.
264. Horwitz, M.S. Transcription regulation *in vitro* by an *E. coli* promoter containing a DNA cruciform in the 3'-5' region. *Nucleic Acids Res.* **1989**, *17*, 5537–5545.
265. Bagga, R.; Ramesh, N.; Brahmachari, S.K. Supercoil-induced unusual structures as transcriptional block. *Nucleic Acids Res.* **1990**, *18*, 3363–3369.
266. Hamada, H.; Bustein, M. Hierarchy of binding sites for chromosomal proteins HMG1 and HMG2 in supercoiled deoxyribonucleic acid. *Biochemistry* **1985**, *24*, 1428–1433.
267. Waga, S.; Mizuno, S.; Yoshida, M. Nonhistone protein HMG1 removes the transcriptional block caused by left-handed Z-form segment in a supercoiled DNA. *Biochem. Biophys. Res. Commun.* **1988**, *153*, 334–339.
268. Sastry, S.S.; Kun, E. The interaction of adenosine diphosphoribosyl transferase (ADPRT) with a cruciform DNA. *Biochem. Biophys. Res. Commun.* **1990**, *167*, 842–847.
269. Oei, S.L.; Herzog, H.; Hirsch-Kauffmann, M.; Schneider, R.; Auer, B.; Schweiger, M. Transcriptional regulation and autoregulation of the human gene for ADP-ribosyltransferase. *Mol. Cell Biochem.* **1994**, *138*, 99–104.

270. Cozzarelli, N.R.; Wang, J.C. *DNA Topology and Its Biological Effects*. Cold Spring Harbor, NY, Cold Spring Harbor Laboratory, 1990.
271. White, J.H.; Bauer, W.R. Superhelical DNA with local substructures. A generalization of the topological constraint in terms of the intersection number and the ladder-like correspondence surface. *J. Mol. Biol.* **1987**, *195*, 205–213.
272. Pearson, C.E.; Ruiz, M.T.; Price, G.B.; Zannis-Hadjopoulos, M. Cruciform DNA binding protein in HeLa cell extracts. *Biochemistry* **1994**, *33*, 14185–14196.
273. Diekmann, S.; Lilley, D.M. The anomalous gel migration of a stable cruciform: temperature and salt dependence, and some comparisons with curved DNA. *Nucleic Acids Res.* **1987**, *15*, 5765–5774.
274. McLean, M.J.; Wells, R.D. The role of DNA sequence in the formation of Z-DNA versus cruciforms in plasmids. *J. Biol. Chem.* **1988**, *263*, 7370–7377.
275. Klysik, J.; Stirdvant, S.; Wells, R. Left-handed DNA. *J. Biol. Chem.* **1982**, *257*, 10152–10158.
276. Burd, J.F.; Wartell, R.M.; Dodgson, J.B.; Wells, R.D. Transmission of stability (telestability) in deoxyribonucleic acid. Physical and enzymatic studies on the duplex block polymer d(C₁₅A₁₅)-d(T₁₅G₁₅). *J. Biol. Chem.* **1975**, *250*, 5109–5113.
277. Klein, R.; Wells, R. Effects of neighboring DNA homopolymers on the biochemical and physical properties of the *Escherichia coli* lactose promoter. *J. Biol. Chem.* **1982**, *257*, 12962–12969.
278. Gierer, A. Model for DNA and protein interactions and the function of the operator. *Nature* **1966**, *212*, 1480–1481.
279. deMassey, B.; Studier, F.W.; Dorgai, L.; Appelbaum, F.; Weisberg, R.A. Enzymes and the sites of genetic recombination: studies with gene-3 endonuclease of phage T7 and with site-affinity mutants of lambda phage. *Cold Spring Harb. Symp. Quant. Biol.* **1984**, *49*, 715–726.
280. West, S.C.; Korner, A. Cleavage of cruciform DNA structures by an activity from *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* **1985**, *82*, 6445–6449.
281. Symington, L.; Kolodner, R. Partial purification of an endonuclease from *Saccharomyces cerevisiae* that cleave Holliday junctions. *Proc. Natl. Acad. Sci. USA* **1985**, *82*, 7247–7251.
282. West, S.C.; Parsons, C.A.; Picksley, S.M. Purification and properties of a nuclease from *Saccharomyces cerevisiae* that cleaves DNA at cruciform junctions. *J. Biol. Chem.* **1987**, *262*, 12752–12758.
283. Taylor, A.F.; Smith, G.R. Action of RecBCD enzyme on cruciform DNA. *J. Mol. Biol.* **1990**, *211*, 117–134.
284. Connolly, B.; Parsons, C.A.; Benson, F.E.; Dunderdale, H.J.; Sharples, G.J.; Lloyd, R.G.; et al. Resolution of Holliday junctions *in vitro* requires the *Escherichia coli* ruvC gene product. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 6063–6067.
285. Parsons, C.A.; Kemper, B.; West, S.C. Interaction of a four-way junction in DNA with T4 endonuclease VII. *J. Biol. Chem.* **1990**, *265*, 9285–9289.
286. Parsons, C.A.; West, S.C. Specificity of binding to four-way junctions in DNA by bacteriophage T7 endonuclease I. *Nucleic Acids Res.* **1990**, *18*, 4377–4384.
287. Bhattacharyya, A.; Murchie, A.I.H.; Bhattacharyya, A.; Clegg, R.M.; Diekmann, S.; von Kitzing, E.; et al. The structure of DNA junctions and their interaction with enzymes. *Eur. J. Biochem.* **1992**, *207*, 285–295.
288. Bianchi, M.E. Interaction of a protein from rat liver nuclei with cruciform DNA. *EMBO J.* **1988**, *7*, 843–849.
289. Bonne-Andrea, C.; Harper, F.; Puvion, E.; Delpech, M. Nuclear accumulation of HMG1 is correlated to DNA synthesis. *Biol. Cell* **1986**, *58*, 185–194.
290. Bustin, M.; Lehn, D.A.; Landsman, D. Structural features of the HMG chromosomal proteins and their genes. *Biochim. Biophys. Acta* **1990**, *1049*, 231–243.

292. Einck, L.; Bustin, M. The intracellular distribution and function of the high mobility group chromosomal proteins. *Exp. Cell Res.* **1985**, *156*, 295–310.
293. Bottger, M.; Vogel, F.; Platzer, U.; Kiessling, U.; Grade, K.; Strauss, M. Condensation of vector DNA by the chromosomal protein HMG1 results in efficient transfection. *Biochim. Biophys. Acta* **1988**, *950*, 221–228.
294. Mathis, D.J.; Kindelis, A.; Spadafora, C. HMG proteins (1+2) from beaded structures when complexed with closed circular DNA. *Nucleic Acids Res.* **1980**, *8*, 2577–2590.
295. Javaherian, K.; Sadeghi, M.; Liu, L.F. Nonhistone proteins HMG1 and HMG2 unwind DNA double helix. *Nucleic Acids Res.* **1979**, *6*, 3569–3580.
296. Javaherian, K.; Liu, L.F.; Wang, J.C. Nonhistone proteins HMG1 and HMG2 change the DNA helical structure. *Science* **1978**, *199*, 1345–1346.
297. Yoshida, M. High glutamic and aspartic regions in nonhistone protein HMB(1+2) unwinds DNA double helical structure. *J. Biochem.* **1987**, *101*, 175–180.
298. Yoshida, M.; Shimura, K. Unwinding of DNA by nonhistone chromosomal protein HMG(1+2) from pig thymus as determined with endonuclease. *J. Biochem.* **1984**, *95*, 117–124.
299. Paull, T.T.; Haykinson, M.; Johnson, R.C. The nonspecific DNA-binding and bending proteins HMG1 and HMG2 promote the assembly of complex nucleoprotein structures. *Genes Dev.* **1993**, *7*, 1521–1534.
300. Pil, P.M.; Chow, C.S.; Lippard, S.J. High-mobility-group 1 protein mediates DNA bending determined by ring closures. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 9465–9469.
301. Hodges-Garcia, Y.; Hagerman, P.J.; Pettijohn, D.E. DNA ring closure mediated by protein HU. *J. Biol. Chem.* **1989**, *264*, 14621–14623.
302. Bianchi, M.E. Prokaryotic HU and eukaryotic HMG1: a kinked relationship. *Mol. Microbiol.* **1994**, *14*, 1–5.
303. Jackson, J.B.; Rill, R.L. Circular dichroism, thermal denaturation, and deoxyribonuclease I digestion studies of nucleosomes highly enriched in high mobility group proteins HMG1 and HMG2. *Biochemistry* **1981**, *20*, 1042–1046.
304. Singh, J.; Dixon, G.H. High mobility group proteins 1 and 2 function as general class II transcription factors. *Biochemistry* **1990**, *29*, 6295–6302.
305. Tremethick, D.J.; Molloy, P.L. Effects of high mobility group proteins 1 and 2 on initiation and elongation of specific transcription by RNA polymerase II *in vitro*. *Nucleic Acids Res.* **1988**, *16*, 11107–11123.
306. Tremethick, D.J.; Molloy, P.L. High mobility group proteins 1 and 2 stimulate transcription *in vitro* by RNA polymerase II and III. *J. Biol. Chem.* **1986**, *261*, 6986–6992.
307. Watt, F.; Molloy, P.L. High mobility group proteins 1 and 2 stimulate binding of a specific transcription factor in the adenovirus major late promoter. *Nucleic Acids Res.* **1988**, *16*, 1471–1486.
308. Bonne, C.; Sautiere, P.; Duguet, M.; De Recondo, A.-M. Identification of a single-stranded DNA binding protein from rat liver with high mobility group protein 1. *J. Biol. Chem.* **1982**, *257*, 2722–2725.
309. Alexandrova, E.A.; Marekov, L.N.; Beltchev, B.G. Involvement of protein HMG1 in DNA replication. *FEBS Lett.* **1984**, *178*, 153–156.
310. Alexandrova, E.A.; Beltchev, B.G. Acetylated HMG1 protein interacts specifically with homologous DNA polymerase alpha *in vitro*. *Biochem. Biophys. Res. Commun.* **1988**, *154*, 918–927.
311. Bianchi, M.E.; Falcioiu, L.; Ferrari, S.; Lilley, D.M.J. The DNA binding site of HMG1 protein is composed of two similar segments (HMG boxes), both of which have counterparts in other eukaryotic regulatory proteins. *EMBO J.* **1992**, *11*, 1055–1063.
312. Ner, S.S. HMGs everywhere. *Curr. Biol.* **1992**, *2*, 208–210.
313. Isackson, P.J.; Fishback, J.L.; Bidney, D.L.; Reeck, G.R. Preferential affinity of high molecular weight high mobility group non-histone chromatin proteins for single-stranded DNA. *J. Biol. Chem.* **1979**, *254*, 5569–5572.
314. Ferrari, S.; Harley, V.R.; Pontigga, A.; Goodfellow, P.N.; Lovell-Badge, R.; Bianchi, M.E. SRY-like HMG1, recognizes sharp angles in DNA. *EMBO J.* **1992**, *11*, 4497–4506.

315. Harley, V.R.; Goodfellow, P.N. The biochemical role of SRY in sex determination. *Mol. Reprod. Dev.* **1994**, *39*, 184–193.
316. Bianchi, M.E.; Beltrame, M.; Paonessa, G. Specific recognition of cruciform DNA by nuclear protein HMG1. *Science* **1989**, *243*, 1056–1059.
317. Gut, S.H.; Bischoff, M.; Hobi, R.; Kuenzle, C.C. Z-DNA binding proteins from bull testis. *Nucleic Acids Res.* **1987**, *15*, 9691–9705.
318. Christen, T.; Bischoff, M.; Hobi, R.; Kuenzle, C.C. High mobility group proteins 1 and 2 bind preferentially to brominated poly(dG-dC)-poly(dG-dC) in the Z-DNA conformation but not to other types of Z-DNA. *FEBS Lett.* **1990**, *267*, 139–141.
319. Rohner, K.J.; Hobi, R.; Kuenzle, C.C. Z-DNA-binding proteins. Identification critically depends on the proper choice of ligands. *J. Biol. Chem.* **1990**, *265*, 19112–19115.
320. Bruhn, S.L.; Pil, P.M.; Essigman, J.M.; Houseman, D.E.; Lippard, S.J. Isolation and characterization of human cDNA clones encoding a high mobility group box protein that recognizes structural distortions to DNA caused by binding of the anticancer agent cisplatin. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 2307–2311.
321. Bianchi, M.E.; Beltrame, M.; Falciola, L. The HMG box motif. In: *Nucleic Acids and Molecular Biology* (Eckstein, F.; Lilley, D.M.J.; Eds.). Springer-Verlag, Berlin, 1992, 112–128.
322. Jantzen, H.-M.; Admon, A.; Bell, S.P.; Tijan, R. Nucleolar transcription factor hUBF contains a DNA-binding motif with homology to HMG proteins. *Nature* **1990**, *344*, 830–836.
323. Travis, A.; Amsterdam, A.; Berlanger, C.; Grosschedl, R. LEF-1 gene encoding a lymphoid-specific protein with an HMG domain, regulates T-cell receptor α enhancer function. *Genes Dev.* **1991**, *5*, 880–894.
324. Diffley, J.F.X.; Stillman, B. Purification of a yeast protein that binds to origins of DNA replication and a transcriptional silencer. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 2120–2124.
325. Shirakata, M.; Huppi, K.; Usuda, S.; Okazaki, K.; Yoshida, K.; Sakano, H. HMG1-related DNA-binding protein isolated with V-(D)-J recombination signal probes. *Mol. Cell Biol.* **1991**, *11*, 4528–4536.
326. Sinclair, A.H.; Berta, P.; Palmer, M.S.; Hawkins, J.R.; Griffiths, B.L.; Smith, M.J.; et al. A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. *Nature* **1990**, *346*, 240–244.
327. Copenhagen, G.P.; Putnam, C.D.; Denton, M.L.; Pikaard, C.S. The RNA polymerase I transcription factor UBF is a sequence-tolerant HMG-box protein that can recognize structured nucleic acids. *Nucleic Acids Res.* **1994**, *22*, 2651–2657.
328. Laudet, V.; Stehelin, D.; Clevers, H. Ancestry and diversity of the HMG box superfamily. *Nucleic Acids Res.* **1993**, *21*, 2493–2501.
329. Varga-Weisz, P.; Zlatanova, J.; Leuba, S.H.; Schroth, G.P.; van Holde, K. Binding of histones H1 and H5 and their globular domains to four-way junction DNA. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 3525–3529.
330. Carballo, M.; Puigdomenech, P.; Palau, J. DNA and histone H1 interact with different domains of HMG1. *EMBO J.* **1983**, *2*, 1759–1764.
331. Kohlstaedt, L.A.; Cole, R.D. Specific interaction between H1 histone and high mobility group protein, HMG1. *Biochemistry* **1994**, *33*, 570–575.
332. Kohlstaedt, L.A.; Sung, E.C.; Fujishige, A.; Cole, R.D. Non-histone chromosomal protein HMG1 modulates the histone H1-induced condensation of DNA. *J. Biol. Chem.* **1987**, *262*, 524–526.
333. Bonne, C.; Harper, C.E.; Sobczak, J.; De Recondo, A.-M. Rat liver HMG1: a physiological nucleosome assembly factor. *EMBO J.* **1984**, *3*, 1193–1199.
334. Zechiedrich, E.L.; Osheroff, N. Eukaryotic topoisomerases recognize nucleic acid topology by preferentially interacting with DNA crossovers. *EMBO J.* **1990**, *9*, 4555–4562.
335. Pognan, F.; Paolletti, C. Does cruciform DNA provide a recognition signal for DNA-topoisomerase II? *Biochimie* **1992**, *74*, 1019–1023.

336. Sekiguchi, J.; Seeman, N.C.; Shuman, S. Resolution of Holliday junctions by eukaryotic DNA topoisomerase. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 785–789.
337. Froelich-Ammon, S.J.; Gale, K.C.; Osherooff, N. Site-specific cleavage of a DNA hairpin by topoisomerase II. *J. Biol. Chem.* **1994**, *269*, 7719–7725.
338. Arndt-Jovin, D.J.; Uduardy, A.; Garner, M.M.; Ritter, S.; Jovin, T.M. Z-DNA binding and inhibition by GTP of *Drosophila* topoisomerase II. *Biochemistry* **1993**, *32*, 4862–4872.
339. McMurray, C.T.; Wilson, W.D.; Douglass, J.O. Hairpin formation within the enhancer region of the human enkephalin gene. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 666–670.
340. Spiro, C.; Richards, J.P.; Chandrasekaran, S.; Brennan, R.G.; McMurray, C.T. Secondary structure creates mismatched base pairs required for high-affinity binding of cAMP response element-binding protein to the human enkephalin enhancer. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 4606–4610.
341. Carmichael, E.P.; Roome, J.M.; Wahl, A.F. Binding of a sequence-specific single-stranded DNA-binding factor to the simian virus 40 core origin inverted repeat domain is cell cycle regulated. *Mol. Cell Biol.* **1993**, *13*, 408–420.
342. Quinn, J.P.; McAllister, J. A preprotachykinin promoter interacts with a sequence specific single stranded DNA binding protein. *Nucleic Acids Res.* **1996**, *21*, 1637–1641.
343. Kim, C.; Snyder, R.O.; Wold, M.S. Binding properties of replication protein A from human and yeast cells. *Mol. Cell Biol.* **1992**, *12*, 3050–3059.
344. Mitsui, Y.; Langridge, R.; Grant, R.C.; Kodama, M.; Wells, R.D.; Shortle, B.E.; et al. Physical and enzymatic studies on poly(dI-dC)-poly(dI-dC), an unusual double-helical DNA. *Nature* **1970**, *228*, 1166–1169.
345. Sutherland, J.C.; Griffen, K.P. Vacuum ultraviolet circular dichroism of poly(dI-dC)-poly(dI-dC): no evidence for a left-handed double helix. *Biopolymers* **1983**, *22*, 1445–1448.
346. Pohl, F.M.; Jovin, T.M. Salt-induced co-operative conformational change of a synthetic DNA: equilibrium and kinetic studies with poly(dG-dC). *J. Mol. Biol.* **1972**, *67*, 375–396.
347. Wang, A.H.J.; Quigley, G.J.; Kolpak, F.J.; Crawford, J.L.; Van Boom, J.H.; Van der Marel, G.; et al. Molecular structure of a left-handed double helical DNA fragment at atomic resolution. *Nature* **1979**, *282*, 680–686.
348. Rich, A.; Nordheim, A.; Wang, A.-J. The chemistry and biology of left-handed Z-DNA. *Annu. Rev. Biochem.* **1984**, *53*, 791–846.
349. Klysik, J.; Zacharias, W.; Galazka, G.; Kwinkowski, M.; Uznanski, B.; Okruszek, A. Structural interconversion of alternating purine-pyrimidine inverted repeats cloned in supercoiled plasmids. *Nucleic Acids Res.* **1988**, *16*, 6915–6933.
350. Nejedly, K.; Klysik, J.; Palecek, E. Supercoil-stabilized left handed DNA in the plasmid (dA-dT)₁₆ insert formed in the presence of Ni²⁺. *FEBS Lett.* **1989**, *243*, 313–317.
351. Singleton, C.K.; Klysik, J.; Stirdvant, S.M.; Wells, R.D. Left-handed Z-DNA is induced by supercoiling in physiological ionic conditions. *Nature* **1982**, *299*, 312–316.
352. Peck, L.J.; Nordheim, A.; Rich, A.; Wang, J.C. Flipping of cloned d(pCpG)_n-d(pGpC)_n DNA sequences from right-to-left-handed helical structure by salt, Co(III), or negative supercoiling. *Proc. Natl. Acad. Sci. USA* **1982**, *79*, 4560–4564.
353. Klysik, J.; Stirdvant, S.M.; Larson, J.E.; Hart, P.A.; Wells, R.D. Left-handed DNA in restriction fragments and a recombinant plasmid. *Nature* **1981**, *290*, 672–677.
354. Stirdvant, S.M.; Klysik, J.; Wells, R.D. Energetic and structural inter-relationship between DNA supercoiling and the right- to left-handed Z helix transitions in recombinant plasmids. *J. Biol. Chem.* **1982**, *257*, 10159–10165.
355. Haniford, D.B.; Pulleyblank, D.E. Facile transition of poly[d(TG)-d(CA)] into a left-handed helix in physiological conditions. *Nature* **1983**, *302*, 632–634.
356. Nordheim, A.; Rich, A. The sequence (dC-dA)_n-(dG-dT)_n forms left-handed Z-DNA in negatively supercoiled plasmids. *Proc. Natl. Acad. Sci. USA* **1983**, *80*, 1821–1825.
357. Ellison, M.J.; Kelleher, R.J., III; Wang, A.H.-J.; Habener, J.F.; Rich, A. Sequence-dependent energetics of the B-Z transition in supercoiled DNA containing nonalternating purine-pyrimidine sequences. *Proc. Natl. Acad. Sci. USA* **1985**, *82*, 8320–8324.

358. Blaho, J.A.; Larson, J.E.; McLean, M.J.; Wells, R.D. Multiple DNA secondary structures in perfect inverted repeat inserts in plasmids. Right-handed B-DNA, cruciforms, and left-handed Z-DNA. *J. Biol. Chem.* **1988**, *263*, 14446–14455.
359. Sinden, R.R.; Kochel, T.J. Reduced 4,5',8-trimethylpsoralen cross-linking of left-handed Z-DNA stabilized by DNA supercoiling. *Biochemistry* **1987**, *26*, 1343–1350.
360. Kochel, T.J.; Sinden, R.R. Analysis of trimethylpsoralen photoreactivity to Z-DNA provides a general *in vivo* assay for Z-DNA: analysis of the hypersensitivity of (GT)_n B-Z junctions. *Bio-techniques* **1988**, *6*, 532–543.
361. Lafer, E.M.; Sousa, R.; Rich, A. Anti-Z-DNA antibody binding can stabilize Z-DNA in relaxed and linear plasmids under physiological conditions. *EMBO J.* **1985**, *4*, 3655–3660.
362. Lafer, E.; Moller, A.; Nordheim, A.; Stollar, B.D.; Rich, A. Antibodies specific for left-handed Z-DNA. *Proc. Natl. Acad. Sci. USA* **1981**, *78*, 3546–3550.
363. Moller, A.; Gabriels, J.E.; Lafer, E.M.; Nordheim, A.; Rich, A.; Stollar, D.B. Monoclonal antibodies recognize different parts of Z-DNA*. *J. Biol. Chem.* **1982**, *257*, 12081–12085.
364. Peck, L.J.; Wang, J.C. Energetics of B-to-Z transition in DNA. *Proc. Natl. Acad. Sci. USA* **1983**, *80*, 6206–6210.
365. Nordheim, A.; Peck, L.J.; Lafer, E.M.; Stollar, B.D.; Wang, J.C.; Rich, A. Supercoiling and left-handed Z-DNA. *Cold Spring Harb. Symp. Quant. Biol.* **1983**, *47 Pt 1*, 93–100.
366. Johnston, B.H.; Rich, A. Chemical probes of DNA conformation: detection of Z-DNA at nucleotide resolution. *Cell* **1985**, *42*, 713–724.
367. McLean, M.J.; Lee, J.W.; Wells, R.D. Characteristics of Z-DNA helices formed by imperfect (purine-pyrimidine) sequences in plasmids. *J. Biol. Chem.* **1988**, *263*, 7378–7385.
368. Rio, P.; Leng, M. N-Hydroxyaminofluorene: a chemical probe for DNA conformation. *J. Mol. Biol.* **1986**, *191*, 569–572.
369. Kochel, T.J.; Sinden, R.R. Hyperreactivity of B-Z junctions to 4,5',8-trimethylpsoralen photo-binding assayed by an exonuclease III/photoreversal mapping procedure. *J. Mol. Biol.* **1989**, *205*, 91–102.
370. Hoepfner, R.W.; Sinden, R.R. Amplified primer extension assay for psoralen photoproducts provides a sensitive assay for a (CG)₆TA(CG)₂(TG)₈Z-DNA torsionally tuned probe: preferential psoralen photobinding to one strand of a B-Z junction. *Biochemistry* **1993**, *32*, 7542–7548.
371. Azorin, F.; Hahn, R.; Rich, A. Restriction endonucleases can be used to study B-Z junction in supercoiled DNA. *Biochemistry* **1984**, *81*, 5714–5718.
372. Singleton, C.K.; Klysik, J.; Wells, R.D. Conformational flexibility of junctions between contiguous B- and Z-DNAs in supercoiled plasmids. *Proc. Natl. Acad. Sci. USA* **1983**, *80*, 2447–2451.
373. Zacharias, W.; Larson, J.E.; Kilpatrick, M.W.; Wells, R.D. HhaI methylase and restriction endonuclease as probes for B to Z DNA conformational changes in d(GCGC) sequences. *Nucleic Acids Res.* **1984**, *12*, 7677–7692.
374. Vardimon, L.; Rich, A. In Z-DNA the sequence G-C-G-C is neither methylated by Hha I methyltransferase nor cleaved by Hha I restriction endonuclease. *Proc. Natl. Acad. Sci. USA* **1984**, *81*, 3268–3272.
375. Zhang, S.; Lockshin, C.; Herbert, A.; Winter, E.; Rich, A. Zootin, a putative Z-DNA binding protein in *Saccharomyces cerevisiae*. *EMBO J.* **1992**, *11*, 3787–3796.
376. Jaworski, A.; Hsieh, W.T.; Blaho, J.A.; Larson, J.E.; Wells, R.D. Left-handed DNA *in vivo*. *Science* **1987**, *238*, 773–777.
- 376a. Haniford, D.B.; Pulleyblank, D.E. The *in vivo* occurrence of Z DNA. *J. Biomol. Struct. Dyn.* **1983**, *1*, 593–609.
377. Rahmouni, A.R.; Wells, R.D. Stabilization of Z DNA *in vivo* by localized supercoiling. *Science* **1989**, *246*, 358–363.
378. Nordheim, A.; Pardue, M.L.; Lafer, E.M.; Moller, A.; Stollar, B.D.; Rich, A. Antibodies to left-handed Z-DNA bind to interband regions of *Drosophila* polytene chromosomes. *Nature* **1981**, *294*, 417–422.

379. Hill, R.J.; Stollar, B.D. Dependence of Z-DNA antibody binding to polytene chromosomes on acid fixation and DNA torsional strain. *Nature* **1983**, *305*, 338–340.
380. Jackson, D.A.; Cook, P.R. A general method for preparing chromatin containing intact DNA. *EMBO J.* **1985**, *4*, 913–918.
381. Santoro, C.; Costanzo, F.; Ciliberto, G. Inhibition of eukaryotic tRNA transcription by potential Z-DNA sequences. *EMBO J.* **1984**, *3*, 1553–1559.
382. Banerjee, R.; Carothers, A.M.; Grunberger, D. Inhibition of the herpes simplex virus thymidine kinase gene transfection in Ltk-cells by potential Z-DNA forming polymers. *Nucleic Acids Res.* **1985**, *13*, 5111–5126.
383. Banerjee, R.; Grunberger, D. Enhanced expression of the bacterial chloramphenicol acetyltransferase gene in mouse cells cotransfected with synthetic polynucleotides able to form Z-DNA. *Proc. Natl. Acad. Sci. USA* **1986**, *83*, 4988–4992.
384. Naylor, L.H.; Clark, E.M. d(TG)_nd(CA)_n sequences upstream of the rat prolactin gene form Z-DNA and inhibit gene transcription. *Nucleic Acids Res.* **1990**, *18*, 1595–1601.
385. Blaho, J.A.; Wells, R.D. Left-handed Z-DNA and genetic recombination. *Prog. Nucleic Acids Res. Mol. Biol.* **1989**, *37*, 107–126.
386. Molineaux, S.M.; Engh, H.; de-Ferra, F.; Hudson, L.; Lazzarini, R.A. Recombination within the myelin basic protein gene created the dysmyelinating shiverer mouse mutation. *Proc. Natl. Acad. Sci. USA* **1986**, *83*, 7542–7546.
387. Weinreb, A.; Collier, D.A.; Birshtein, B.K.; Wells, R.D. Left-handed Z-DNA and intramolecular triplex formation at the site of an unequal sister chromatid exchange. *J. Biol. Chem.* **1990**, *265*, 1352–1359.
388. Weinreb, A.; Katzenberg, D.R.; Gilmore, G.L.; Birshtein, B.K. Site of unequal sister chromatid exchange contains a potential Z-DNA-forming tract. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 529–533.
389. Boehm, T.; Mengle-Gaw, L.; Kees, U.R.; Spurr, N.; Lavenir, I.; Forster, A.; et al. Alternating purine-pyrimidine tracts may promote chromosomal translocations seen in a variety of human lymphoid tumours. *EMBO J.* **1989**, *8*, 2621–2631.
390. Wahls, W.P.; Wallace, L.J.; Moore, P.D. The Z-DNA motif d(TG)₃₀ promotes reception of information during gene conversion events while stimulating homologous recombination in human cells in culture. *Mol. Cell Biol.* **1990**, *10*, 785–793.
391. Choo, K.B.; Lee, H.H.; Liew, L.N.; Chong, K.Y.; Chou, H.F. Analysis of the unoccupied site of an integrated human papillomavirus 16 sequence in a cervical carcinoma. *Virology* **1990**, *178*, 621–625.
392. Kmiec, E.B.; Angelides, K.J.; Holloman, W.K. Left-handed DNA and the synaptic pairing reaction promoted by *Ustilago* rec1 protein. *Cell* **1985**, *40*, 139–145.
393. Blaho, J.A.; Wells, R.D. Left-handed Z-DNA binding by the recA protein of *Escherichia coli*. *J. Biol. Chem.* **1987**, *262*, 6082–6088 [published erratum appears in *J. Biol. Chem.* **1988**, *263*, 11015].
394. Fishel, R.A.; Detmer, K.; Rich, A. Identification of homologous pairing and strand-exchange activity from a human tumor cell line based on Z-DNA affinity chromatography. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 36–40.
395. Herbert, A.G.; Spitzner, J.R.; Lowenhaupt, K.; Rich, A. Z-DNA binding protein from chicken blood nuclei. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 3339–3342.
396. Gut, S.H.; Bischoff, M.; Hobi, R.; Kuenzle, C.C. Z-DNA-binding proteins from bull testis. *Nucleic Acids Res.* **1987**, *15*, 9691–9705.
397. Krishna, P.; Kennedy, B.P.; van de Sande, J.H.; McGhee, J.D. Yolk proteins from nematodes, chickens, and frogs bind strongly and preferentially to left-handed Z-DNA. *J. Biol. Chem.* **1988**, *263*, 19066–19070.
398. Lafer, E.M.; Sousa, R.J.; Rich, A. Z-DNA-binding proteins in *Escherichia coli* purification, generation of monoclonal antibodies and gene isolation. *J. Mol. Biol.* **1988**, *203*, 511–516.

399. Lafer, E.M.; Sousa, R.; Rosen, B.; Hsu, A.; Rich, A. Isolation and characterization of Z-DNA binding proteins from wheat germ. *Biochemistry* **1985**, *24*, 5070–5076.
400. Nordheim, A.; Lafer, E.M.; Peck, L.J.; Wang, J.C.; Stollar, B.D.; Rich, A. Negatively supercoiled plasmids contain left-handed Z-DNA segments as detected by specific antibody binding. *Cell* **1982**, *31*, 309–318.
401. Feisenfeld, G.; Davies, D.R.; Rich, A. Formation of a three-stranded polynucleotide molecule. *J. Am. Chem. Soc.* **1957**, *79*, 2023–2024.
402. Wells, R.D.; Collier, D.A.; Hanvey, J.C.; Shimizu, M.; Wohlrab, F. The chemistry and biology of unusual DNA structures adopted by oligopurine-oligopyrimidine sequences. *FASEB J.* **1988**, *2*, 2939–2949.
403. Thuong, N.T.; Hélène, C. Sequence-specific recognition and modification of double-helical DNA by oligonucleotides. *Angew. Chem. Int. Ed. Engl.* **1993**, *32*, 666–690.
404. Mirkin, S.M.; Frank-Kamenetskii, M.D. H-DNA and related structures. *Annu. Rev. Biophys. Biomol. Struct.* **1994**, *23*, 541–576.
405. Radhakrishnan, I.; Patel, D.J. DNA tripleplexes: solution structures, hydration sites, energetics, interactions, and function. *Biochemistry* **1994**, *33*, 11405–11416.
406. Frank-Kamenetskii, M.D.; Mirkin, S.M. Triplex DNA structures. *Annu. Rev. Biochem.* **1995**, *64*, 65–95.
407. Chubb, J.M.; Hogan, M.E. Human therapeutics based on triple helix technology. *Trends Biotechnol.* **1992**, *10*, 132–136.
408. Helene, C.; Thuong, N.T.; Harrel-Bellan, A. Control of gene expression by triple helix-forming oligonucleotides. The antigene strategy. *Ann. N.Y. Acad. Sci.* **1992**, *660*, 27–36.
409. Strobel, S.A.; Dervan, P.B. Triple helix-mediated single-site enzymatic cleavage of megabase genomic DNA. *Methods Enzymol.* **1992**, *216*, 309–321.
410. Maher, L.J., III. Prospects for the therapeutic use of antigenic oligonucleotides. *Cancer Invest.* **1996**, *14*, 66–82.
411. Felsenfeld, G.; Miles, H.T. The physical and chemical properties of nucleic acids. *Annu. Rev. Biochem.* **1967**, *26*, 407–468.
412. Michelson, A.M.; Massoulie, J.; Guschlbauer, W. Synthetic polynucleotides. *Prog. Nucleic Acids Res. Mol. Biol.* **1967**, *6*, 83–141.
413. Cheng, Y.K.; Pettitt, B.M. Stabilities of double- and triple-strand helical nucleic acids. *Prog. Biophys. Mol. Biol.* **1992**, *58*, 225–257.
414. Moser, H.E.; Dervan, P.B. Sequence-specific cleavage of double helical DNA by triple helix formation. *Science* **1987**, *238*, 645–650.
415. Le Doan, T.; Perrouault, L.; Praseuth, D.; Habhou, N.; Decoult, J.L.; Thuong, N.T.; et al. Sequence-specific recognition, photocrosslinking and cleavage of the DNA double-helix by an oligo-[α]-thymidilate covalently linked to an azidoproflavine derivative. *Nucleic Acids Res.* **1987**, *15*, 7749–7760.
416. Cooney, M.; Czernuszewicz, G.; Postel, E.H.; Flint, S.J.; Hogan, M.E. Site-specific oligonucleotide binding represses transcription of the human *c-myc* gene *in vitro*. *Science* **1988**, *241*, 456–459.
417. Lyamichev, V.I.; Mirkin, S.M.; Frank-Kamenetskii, M.D. Structures of homopurine-homopyrimidine tract in superhelical DNA. *J. Biomol. Struct. Dyn.* **1986**, *3*, 667–669.
418. Kohwi, Y.; Kohwi-Shigematsu, T. Magnesium ion-dependent triple-helix structure formed by homopurine-homopyrimidine sequences in supercoiled plasmid DNA. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 3781–3785.
419. Piriou, J.M.; Ketterle, C.; Gabarro-Arpa, J.; Cognet, J.A.H.; Le Bret, M. A database of 32 DNA triplets to study triple helices by molecular mechanics and dynamics. *Biophys. Chem.* **1994**, *50*, 323–343.
420. van Vlijmen, H.W.T.; Ramé, G.L.; Pettitt, B.M. A study of model energetics and conformational properties of polynucleotide tripleplexes. *Biopolymers* **1990**, *30*, 517–532.

421. Helene, C.; Lancelot, G. Interactions between functional groups in protein-nucleic acid associations. *Prog. Biophys. Mol. Biol.* **1982**, *39*, 1–68.
422. Laughton, C.A.; Neidle, S. Prediction of the structure of the Y+R-R(+) type DNA triple helix by molecular modelling. *Nucleic Acids Res.* **1992**, *20*, 6535–6541.
423. Arnott, S.; Selsing, E. Structures for the polynucleotide complexes poly(dA)*poly(dT) and poly(dT)*poly(dA)*poly(dT). *J. Mol. Biol.* **1974**, *88*, 509–521.
424. Raghunathan, G.; Miles, H.T.; Sasisekharan, V. Symmetry and molecular structure of a DNA triple helix: d(T)_n·d(A)_n·d(T)_n. *Biochemistry* **1993**, *32*, 455–462.
425. Macaya, R.; Wang, E.; Schultze, P.; Sklenar, V.; Feigson, J. Proton nuclear magnetic resonance assignments and structural characterization of an intramolecular DNA triplex. *J. Mol. Biol.* **1992**, *225*, 755–773.
426. Shin, C.; Koo, H.S. Helical periodicity of GA-alternating triple-stranded DNA. *Biochemistry* **1996**, *35*, 968–972.
427. Htun, H.; Dahlberg, J.E. Topology and formation of triple-stranded H-DNA. *Science* **1989**, *243*, 1571–1576.
428. Hanvey, J.C.; Shimizu, M.; Wells, R.D. Intramolecular DNA triplexes in supercoiled plasmids. II. Effect of base composition and noncentral interruptions on formation and stability. *J. Biol. Chem.* **1989**, *264*, 5950–5956.
429. Panyutin, I.G.; Wells, R.D. Nodule DNA in the (GA)₃₇(CT)₃₇ insert in superhelical plasmids. *J. Biol. Chem.* **1992**, *267*, 5495–5501.
430. Maher, L.J.; Dervan, P.B.; Wold, B.J. Kinetic analysis of oligodeoxyribonucleotide-directed triple-helix formation on DNA. *Biochemistry* **1990**, *29*, 8820–8826.
431. Rougee, M.; Faucon, B.; Mergny, J.L.; Barcelo, F.; Giovannangeli, C.; Garestier, T.; et al. Kinetics and thermodynamics of triple-helix formation: effects of ionic strength and mismatches. *Biochemistry* **1992**, *31*, 9269–9278.
432. Durland, R.H.; Kessler, D.J.; Gunnell, S.; Duvic, M.; Pettitt, B.M.; Hogan, M.E. Binding of triple helix forming oligonucleotides to sites in gene promoters. *Biochemistry* **1991**, *30*, 9246–9255.
433. Krakauer, H.; Sturtevant, J.M. Heats of the helix-coil transitions of the poly A-poly U complexes. *Biopolymers* **1968**, *6*, 491–512.
434. Glaser, R.; Gabay, E.J. Topography of nucleic acid helices in solutions. III. Interactions of spermine and spermidine derivatives with polyadenylic-polyuridylic and polyinosinic-polycytidyl acid helices. *Biopolymers* **1968**, *6*, 243–254.
435. Hampel, K.J.; Crosson, P.; Lee, J.S. Polyamines favor DNA triplex formation at neutral pH. *Biochemistry* **1991**, *30*, 4455–4459.
436. Hanvey, J.C.; Williams, E.M.; Besterman, J.M. DNA triple-helix formation at physiologic pH and temperature. *Antisense Res. Dev.* **1991**, *1*, 307–317.
437. Singleton, S.F.; Dervan, P.B. Equilibrium association constants for oligonucleotide-directed triple helix formation at single DNA sites: linkage to cation valence and concentration. *Biochemistry* **1993**, *32*, 13171–13179.
438. Thomas, T.; Thomas, T.J. Selectivity of polyamines in triplex DNA stabilization. *Biochemistry* **1993**, *32*, 14068–14074.
439. Potaman, V.N.; Sinden, R.R. Stabilization of triple-helical nucleic acids by basic oligopeptides. *Biochemistry* **1995**, *34*, 14885–14992.
440. Kiessling, L.L.; Griffin, L.C.; Dervan, P.B. Flanking sequence effects within the pyrimidine triple-helix motif characterized by affinity cleaving. *Biochemistry* **1992**, *31*, 2829–2834.
441. Völker, J.; Klump, H.H. Electrostatic effects in DNA triple helices. *Biochemistry* **1994**, *33*, 13502–13508.
442. Darnell, J.; Lodish, H.; Baltimore, D. *Molecular Cell Biology*. Scientific American Books, New York, 1990.
- 442a. Sarhan, S.; Seiler, N. On the subcellular localization of polyamines. *Biol. Chem. Hoppe-Seyler* **1989**, *370*, 1279–1284.

443. Davis, R.W.; Morris, D.R.; Coffino, P. Sequestered end products and enzyme regulation: the case of ornitine decarboxylase. *Microbiol. Rev.* **1992**, *56*, 280–290.
444. Frank-Kamenetskii, M.D. Protonated DNA structures. *Methods Enzymol.* **1992**, *211*, 180–191.
445. Bernues, J.; Beltran, R.; Casasnovas, J.M.; Azorin, F. DNA-sequence and metal-ion specificity of the formation of *H-DNA. *Nucleic Acids Res.* **1990**, *18*, 4067–4073.
446. Roberts, R.W.; Crothers, D.M. Specificity and stringency in DNA triplex formation. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 9397–9401.
447. Weerasinghe, S.; Smith, P.E.; Mohan, V.; Cheng, Y.K.; Pettitt, B.M. Nanosecond dynamics and structure of a model DNA triple helix in saltwater solution. *J. Am. Chem. Soc.* **1995**, *117*, 2147–2158.
448. Lyamichev, V.I.; Mirkin, S.M.; Frank-Kamenetskii, M.D. A pH-dependent structural transition in the homopurine-homopyrimidine tract in superhelical DNA. *J. Biomol. Struct. Dyn.* **1985**, *3*, 327–338.
449. Lyamichev, V.I.; Mirkin, S.M.; Kumarev, V.P.; Baranova, L.V.; Vologodskii, A.V.; Frank-Kamenetskii, M.D. Energetics of the B-H transition in supercoiled DNA carrying d(CT)_xd(AG)_x and d(C)_nd(G)_n inserts. *Nucleic Acids Res.* **1989**, *17*, 9417.
450. Schroth, G.P.; Ho, P.S. Occurrence of potential cruciform and H-DNA forming sequences in genomic DNA. *Nucleic Acids Res.* **1995**, *23*, 1977–1983.
451. Horne, D.A.; Dervan, P.B. Recognition of mixed-sequence duplex DNA by alternate-strand triple-helix formation. *J. Am. Chem. Soc.* **1990**, *112*, 2435–2437.
- 451a. Behe, M.J. An overabundance of long oligopurine tracts occurs in the genome of simple and complex eukaryotes. *Nucleic Acids Res.* **1995**, *23*, 507–511.
452. Pilch, D.S.; Breslauer, K.J. Ligand-induced formation of nucleic acid triple helices. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 9332–9336.
453. Duval-Valentin, G.; de Bizemont, T.; Takasugi, M.; Mergny, J.L.; Bisagni, E.; Helene, C. Triple-helix specific ligands stabilize H-DNA conformation. *J. Mol. Biol.* **1995**, *247*, 847–858.
454. Wilson, W.D.; Mizan, S.; Tanious, F.A.; Yao, S.; Zon, G. The interaction of intercalators and groove-binding agents with DNA triple-helical structures: the influence of ligand structure, DNA backbone modifications and sequence. *J. Mol. Recogn.* **1994**, *7*, 89–98.
455. Nielsen, P.E.; Egholm, M.; Buchardt, O. Peptide nucleic acid (PNA). A DNA mimic with a peptide backbone. *Bioconjug. Chem.* **1994**, *5*, 3–7.
456. Malkov, V.A.; Voloshin, O.N.; Soyfer, V.N.; Frank-Kamenetskii, M.D. Cation and sequence effects on stability of intermolecular pyrimidine-purine-purine triplex. *Nucleic Acids Res.* **1993**, *21*, 585–591.
457. Belotserkovskii, B.P.; Veselkov, A.G.; Filippov, S.A.; Dobrynnin, V.N.; Mirkin, S.M.; Frank-Kamenetskii, M.D. Formation of intramolecular triplex in homopurine-homopyrimidine mirror repeats with point substitutions. *Nucleic Acids Res.* **1990**, *18*, 6621–6624.
458. Xodo, L.E.; Alunni-Fabbroni, M.; Manzini, G.; Quadrifoglio, F. Sequence-specific DNA-triplex formation at imperfect homopurine-homopyrimidine sequences within a DNA plasmid. *Eur. J. Biochem.* **1993**, *212*, 395–401.
459. Wells, R.D.; Amirhaeri, S.; Blaho, J.A.; Collier, D.A.; Dohrman, A.; Griffin, J.A.; et al. Biology and chemistry of Z-DNA and tripleplexes. In: *The Bacterial Chromosome* (Drlka, K.; Riley, M.; Eds.). American Society for Microbiology, Washington, DC, 1990, pp. 187–194.
460. Collier, D.A.; Wells, R.D. Effect of length, supercoiling, and pH on intramolecular triplex formation. Multiple conformers at pur-pyr mirror repeats. *J. Biol. Chem.* **1990**, *265*, 10652–10658.
461. Kohwi-Shigematsu, T.; Kohwi, Y. Detection of triple-helix related structures adopted by poly(dG)-poly(dC) sequences in supercoiled plasmid DNA. *Nucleic Acids Res.* **1991**, *19*, 4267–4271.
462. Naylor, R.; Gilham, P.T. Studies on some interactions and reactions of oligonucleotides in aqueous solution. *Biochemistry* **1966**, *5*, 2722–2728.

463. Cassani, G.R.; Bollum, F.J. Oligodeoxythymidilate:polydeoxyadenylate and oligodeoxyadenylate: polydeoxythymidilate interactions. *Biochemistry* **1969**, *8*, 3928–3936.
464. Broitman, S.L.; Im, D.D.; Fresco, J.R. Formation of the triple-stranded polynucleotide helix, poly(A·U·U). *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 5120–5124.
465. Cheng, A.J.; Van Dyke, M.W. Oligodeoxyribonucleotide length and sequence effects on intermolecular purine-purine-pyrimidine triple-helix formation. *Nucleic Acids Res.* **1994**, *22*, 4742–4747.
466. Johnston, B.H. The S1-sensitive form of $d(C-T)_n \cdot d(A-G)_n$: chemical evidence for a three-stranded structure in plasmids. *Science* **1988**, *241*, 1800–1804.
467. Hanvey, J.C.; Klysik, J.; Wells, R.D. Influence of DNA sequence on the formation of non-B right-handed helices in oligopurine-oligopyrimidine inserts in plasmids. *J. Biol. Chem.* **1988**, *263*, 7386–7396.
468. Lyamichev, V.I.; Frank-Kamenetskii, M.D.; Soyfer, V.N. Protection against UV-induced pyrimidine dimerization in DNA by triplex formation. *Nature* **1990**, *344*, 568–570.
469. Ussery, D.W.; Sinden, R.R. Environmental influences on the *in vivo* level of intramolecular triplex DNA in *Escherichia coli*. *Biochemistry* **1993**, *32*, 6206–6213.
470. Lee, J.S.; Burkholder, G.D.; Latimer, L.J.P.; Haug, B.L.; Braun, R.P. A monoclonal antibody to triplex DNA binds to eucaryotic chromosomes. *Nucleic Acids Res.* **1987**, *15*, 1047–1061.
471. Agazie, Y.M.; Lee, J.S.; Burkholder, G.D. Characterization of new monoclonal antibody to triplex DNA and immunofluorescent staining of mammalian chromosomes. *J. Biol. Chem.* **1994**, *269*, 7019–7023.
472. Lee, J.S.; Latimer, L.J.; Haug, B.L.; Pulleyblank, D.E.; Skinner, D.M.; Burkholder, G.D. Triplex DNA in plasmids and chromosomes. *Gene* **1989**, *82*, 191–199.
473. Karlovsky, P.; Pećinka, P.; Vojtisková, M.; Makaturova, E.; Paleček, E. Protonated triplex DNA in *E. coli* cells as detected by chemical probing. *FEBS Lett.* **1990**, *274*, 39–42.
474. Kohwi, Y.; Makhosyan, S.R.; Kohwi-Shigematsu, T. Intramolecular dG-dG-dC triplex detected in *Escherichia coli* cells. *J. Mol. Biol.* **1992**, *223*, 817–822.
475. Glaser, R.L.; Thomas, G.H.; Siegfried, E.; Elgin, S.C.; Lis, J.T. Optimal heat-induced expression of the *Drosophila hsp26* gene requires a promoter sequence containing $(CT)_n \cdot (GA)_n$ repeats. *J. Mol. Biol.* **1990**, *211*, 751–761.
476. Parniewski, P.; Kwinkowski, M.; Wilk, A.; Klysik, J. Dam methyltransferase sites located within the loop region of the oligopurine-oligopyrimidine sequences capable of forming H-DNA are undermethylated *in vivo*. *Nucleic Acids Res.* **1990**, *18*, 605–611.
477. Jaworski, A.; Blaho, J.A.; Larson, J.E.; Shimizu, M.; Wells, R.D. Tetracycline promoter mutations decrease non-B DNA structural transitions, negative linking differences and deletions in recombinant plasmids in *Escherichia coli*. *J. Mol. Biol.* **1989**, *207*, 513–526.
478. Sarkar, P.S.; Brahmachari, S.K. Intramolecular triplex potential sequence within a gene down regulates its expression *in vivo*. *Nucleic Acids Res.* **1992**, *20*, 5713–5718.
479. Rao, B.S. Pausing of simian virus 40 DNA replication fork movement *in vivo* by $(dG-dA)_n \cdot (dT-dC)_n$ tracts. *Gene* **1994**, *140*, 233–237.
480. Kinniburgh, A.J.; Firulli, A.B.; Kolluri, R. DNA triplexes and regulation of the *c-myc* gene. *Gene* **1994**, *149*, 93–100.
481. Droege, P. Transcription-driven site-specific DNA recombination *in vitro*. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 2759–2763.
482. Bowater, R.P.; Chen, D.; Lilley, D.M.J. Elevated unconstrained supercoiling of plasmid DNA generated by transcription and translation of the tetracycline resistance gene in eubacteria. *Biochemistry* **1994**, *33*, 9266–9275.
483. van Holde, K.; Zlatanova, J. Unusual DNA structures, chromatin and transcription. *Bioessays* **1994**, *16*, 59–68.
484. Lee, J.S.; Woodsworth, M.L.; Latimer, J.P.; Morgan, A.R. Poly(pyrimidine)-poly(purine) synthetic DNAs containing 5-methylcytosine form stable triplexes at neutral pH. *Nucleic Acids Res.* **1984**, *12*, 6603–6614.

485. Gee, J.E.; Blume, S.; Snyder, R.C.; Ray, R.; Miller, D.M. Triplex formation prevents Sp1 binding to the dihydrofolate reductase promoter. *J. Biol. Chem.* **1992**, *267*, 11163–11167.
486. Postel, E.H.; Berberich, S.J.; Flint, S.J.; Ferrone, C.A. Human c-myc transcription factor PuF identified as nm23-H2 nucleoside diphosphate kinase, a candidate suppressor of tumor metastasis. *Science* **1993**, *261*, 478–480.
487. Kolluri, R.; Torey, T.A.; Kinniburgh, A.J. A CT promoter element binding protein: definition of a double-strand and a novel single-strand DNA-binding motif. *Nucleic Acids Res.* **1992**, *20*, 111–116.
488. Yee, H.A.; Wong, A.K.C.; van de Sande, J.H.; Rattner, J.B. Identification of novel single-stranded d(TC)_n binding proteins in several mammalian species. *Nucleic Acids Res.* **1991**, *19*, 949–953.
489. Muraiso, T.; Nomoto, S.; Yamazaki, H.; Mishima, Y.; Kominami, R. A single-stranded DNA binding protein from mouse tumor cells specifically recognizes the C-rich strand of the (AGG:CCT)_n repeats that can alter DNA conformation. *Nucleic Acids Res.* **1992**, *20*, 6631–6635.
490. Goller, M.; Funke, B.; Gehe-Becker, C.; Kroger, B.; Lottspeich, F.; Horak, I. Murine protein which binds preferentially to oligo-C-rich single-stranded nucleic acids. *Nucleic Acids Res.* **1994**, *22*, 1885–1889.
491. Aharoni, A.; Baran, N.; Manor, H. Characterization of a multisubunit human protein which selectively binds single stranded d(GA)_n and d(TC)_n sequence repeats in DNA. *Nucleic Acids Res.* **1993**, *21*, 5221–5228.
492. Hollingworth, M.A.; Closken, C.; Harris, A.; McDonald, C.D.; Pahwa, G.S.; Maher, L.J., III. A nuclear factor that binds purine-rich, single-stranded oligonucleotides derived from S1-sensitive elements upstream of the CFTR gene and the MUC1 gene. *Nucleic Acids Res.* **1994**, *22*, 1138–1146.
493. Kiyama, R.; Camerini-Otero, R.D. A triplex DNA-binding protein from human cells: Purification and characterization. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 10450–10454.
494. Chen, A.; Reyes, A.; Akeson, R. A homopurine:homopyrimidine sequence derived from the rat neuronal cell adhesion molecule-encoding gene alters expression in transient transfections. *Gene* **1993**, *128*, 211–218.
495. Raghu, G.; Tevosian, S.; Anant, S.; Subramanian, K.N.; George, D.L.; Mirkin, S.M. Transcriptional activity of the homopurine-homopyrimidine repeat of the c-Ki-ras promoter is independent of its H-forming potential. *Nucleic Acids Res.* **1994**, *22*, 3271–3279.
496. Bucher, P.; Yagil, G. Occurrence of oligopurine-oligopyrimidine tracts in eukaryotic and prokaryotic genes. *DNA Seq.* **1991**, *1*, 157–172.
497. Tripathi, J.; Brahmachari, S.K. Distribution of simple repetitive(TG/CA)_n and (CT/AG)_n sequences in human and rodent genomes. *J. Biomol. Struct. Dyn.* **1991**, *9*, 387–397.
498. Usdin, K.; Furano, A.V. Insertion of L1 elements into sites that can form non-B DNA. Interactions of non-B DNA-forming sequences. *J. Biol. Chem.* **1989**, *264*, 20736–20743.
499. Pestov, D.G.; Dayn, A.; Siyanova, E.Y.; George, D.L.; Mirkin, S.M. H-DNA and Z-DNA in the mouse c-Ki-ras promoter. *Nucleic Acids Res.* **1991**, *19*, 6527–6532.
500. Nelson, K.L.; Becker, N.A.; Pahwa, G.S.; Hollingworth, M.A.; Maher, L.J., III. Potential for H-DNA in the human MUC1 mucin gene promoter. *J. Biol. Chem.* **1996**, *271*, 18061–18067.
501. Potaman, V.N.; Ussery, D.W.; Sinden, R.R. The formation of a combined H-DNA/open TATA box structure in the promoter sequence of the human Na⁺, K-ATPase α 2 gene. *J. Biol. Chem.* **1996**, *271*, 13441–13447.
502. Menzel, R.; Gellert, M. Regulation of the genes for *E. coli* DNA gyrase: homeostatic control of DNA supercoiling. *Cell* **1983**, *34*, 105–113.
503. Borowiec, J.A.; Gralla, J.D. Supercoiling response of the lac ps promoter *in vitro*. *J. Mol. Biol.* **1985**, *184*, 587–598.
504. Brahms, J.G.; Dargouge, O.; Brahms, S.; Ohara, Y.; Vagner, V. Activation and inhibition of transcription by supercoiling. *J. Mol. Biol.* **1985**, *181*, 455–465.

505. Dorman, C.J.; Barr, G.C.; Bhria, N.N.; Higgins, C.F. DNA supercoiling and the anaerobic and growth phase regulation of tonB gene expression. *J. Bacteriol.* **1988**, *170*, 2816–2826.
506. Parvin, J.D.; Sharp, P.A. DNA topology and a minimal set of basal factors for transcription by RNA polymerase II. *Cell* **1993**, *73*, 533–540.
507. Kato, M.; Shimizu, N. Effect of the potential triplex DNA region on the *in vivo* expression of bacterial β -lactamase gene in superhelical recombinant plasmids. *J. Biochem.* **1992**, *112*, 492–494.
508. Kohwi, Y.; Kohwi-Shigematsu, T. Altered gene expression correlates with DNA structure. *Genes Dev.* **1991**, *5*, 2547–2554.
509. Clark, S.P.; Lewis, C.D.; Felsenfeld, G. Properties of BPG1, a poly(dG)-binding protein from chicken erythrocytes. *Nucleic Acids Res.* **1990**, *18*, 5119–5126.
510. Frank-Kamenetskii, M.D.; Malkov, V.A.; Voloshin, O.N.; Soyfer, V.N. Stabilization of PyPuPu triplets with bivalent cations. *Nucleic Acids Symp. Ser.* **1991**, *24*, 159–162.
511. Johnson, C.A.; Jinno, Y.; Merlino, G.T. Modulation of epidermal growth factor receptor proto-oncogene transcription by a promoter site sensitive to S1 nuclease. *Mol. Cell. Biol.* **1988**, *8*, 4174–4184.
512. Davis, T.L.; Firulli, A.B.; Kinniburgh, A.J. Ribonucleoprotein and protein factors bind to an H-DNA-forming c-myc DNA element: possible regulators of the c-myc gene. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 9682–9686.
513. Postel, E.H.; Mango, S.E.; Flint, S.J. A nuclease-hypersensitive element of the human c-myc promoter interacts with a transcription initiation factor. *Mol. Cell. Biol.* **1989**, *9*, 5123–5133.
514. Mavrothalassitis, G.J.; Watson, D.K.; Papas, T.S. Molecular and functional characterization of the promoter of ETS2, the human c-ets-2 gene. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 1047–1051.
515. Santra, M.; Danielson, K.G.; Iozzo, R.V. Structural and functional characterization of the human decorin gene promoter. A homopurine-homopyrimidine S1 nuclease-sensitive region is involved in transcriptional control. *J. Biol. Chem.* **1994**, *269*, 579–587.
516. Hoffman, E.K.; Trusko, S.P.; Murphy, M.; George, D.L. An S1 nuclease-sensitive homopurine/homopyrimidine domain in the c-Ki-ras promoter interacts with a nuclear factor. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 2705–2709.
517. Lafyatis, R.; Denhez, F.; Williams, T.; Sporn, M.; Roberts, A. Sequence-specific protein binding to and activation of the TGF- β 3 promoter through a repeated TCCC motif. *Nucleic Acids Res.* **1991**, *19*, 6419–6425.
518. Chung, Y.T.; Keller, B. Regulatory elements mediating transcription from the *Drosophila melanogaster* actin 5C proximal promoter. *Mol. Cell. Biol.* **1990**, *10*, 206–216.
519. Firulli, A.B.; Maibenco, D.C.; Kinniburgh, A.J. Triplex forming ability of a c-myc promoter element predicts promoter strength. *Arch. Biochem. Biophys.* **1994**, *310*, 236–242.
520. Mollegaard, N.E.; Buchardt, O.; Egholm, M.; Nielsen, P.E. Peptide nucleic acid-DNA strand displacement loops as artificial transcription promoters. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 3892–3895.
521. Daube, S.S.; von Hippel, P.H. Functional transcription elongation complexes from synthetic RNA-DNA bubble duplexes. *Science* **1992**, *258*, 1320–1324.
522. Biggin, M.D.; Tjian, R. Transcription factors that activate the Ultrabithorax promoter in developmentally staged extracts. *Cell* **1988**, *53*, 699.
523. Grigoriev, M.; Praseuth, D.; Robin, P.; Hemar, A.; Saison-Behmoaras, T.; Dautry-Varsat, A.; et al. A triple helix-forming oligonucleotide-intercalator conjugate acts as a transcriptional repressor via inhibition of NF κ B binding to interleukin-2 receptor α -regulatory sequence. *J. Biol. Chem.* **1992**, *267*, 3389–3395.
524. Spencer, C.A.; Groudine, M. Transcription elongation and eukaryotic gene regulation. *Oncogene* **1990**, *5*, 777–785.
525. Ulrich, M.J.; Gray, W.J.; Ley, T.J. An intramolecular DNA triplex is disrupted by point mutations associated with hereditary persistence of fetal hemoglobin. *J. Biol. Chem.* **1992**, *267*, 18649–18658.

526. Rao, B.S.; Manor, H.; Martin, R.G. Pausing in simian virus 40 DNA replication by a sequence containing (dG-dA)₂₇-(dT-dC)₂₇. *Nucleic Acids Res.* **1988**, *16*, 8077–8094.
527. Baran, N.; Lapidot, A.; Manor, H. Formation of DNA triplexes accounts for arrests of DNA synthesis at d(TC)_n and d(GA)_n tracts. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 507–511.
528. Lapidot, A.; Baran, N.; Manor, H. (dT-dC)_n and (dG-dA)_n tracts arrest single stranded DNA replication *in vitro*. *Nucleic Acids Res.* **1989**, *17*, 883–900.
529. Samadashwily, G.M.; Dayn, A.; Mirkin, S. Suicidal nucleotide sequences for DNA polymerization. *EMBO J.* **1993**, *12*, 4975–4983.
530. Samadashwily, G.M.; Mirkin, S.M. Trapping DNA polymerases using triplex-forming oligodeoxyribonucleotides. *Gene* **1994**, *149*, 127–136.
531. Maine, I.P.; Kodadek, T. Efficient unwinding of triplex DNA by a DNA helicase. *Biochem. Biophys. Res. Commun.* **1994**, *204*, 1119–1124.
532. Kopel, V.; Pozner, A.; Baran, N.; Manor, H. Unwinding of the third strand of a DNA triple helix, a novel activity of the SV40 large T-antigen helicase. *Nucleic Acids Res.* **1996**, *24*, 330–335.
533. Baran, N.; Neer, A.; Manor, H. “Onion skin” replication of integrated polyoma virus DNA and flanking sequences in polyoma-transformed rat cells: termination within a specific cellular DNA segment. *Proc. Natl. Acad. Sci. USA* **1983**, *80*, 105–109.
534. Brinton, B.T.; Caddle, M.S.; Heintz, N.H. Position and orientation-dependent effects of a eukaryotic Z-triplex DNA motif on episomal DNA replication in COS-7 cells. *J. Biol. Chem.* **1991**, *266*, 5153–5161.
535. Birnboim, H.C. Spacing of polypyrimidine regions in mouse DNA as determined by poly(adenylate,guanylate) binding. *J. Mol. Biol.* **1978**, *121*, 541–559.
536. Manor, H.; Rao, B.S.; Martin, R.G. Abundance and degree of dispersion of genomic d(CA)_n-d(TC)_n sequences. *J. Mol. Evol.* **1988**, *27*, 96–101.
537. Lee, J.S.; Morgan, A.R. Novel aspects of the structure of the *Escherichia coli* nucleoid investigated by a rapid sedimentation assay. *Can. J. Biochem.* **1982**, *60*, 952–961.
538. Hampel, K.J.; Burkholder, G.D.; Lee, J.S. Plasmid dimerization mediated by triplex formation between polypyrimidine-polypurine repeats. *Biochemistry* **1993**, *32*, 1072–1077.
539. Blackburn, E.H. Telomeres. *Trends Biol. Sci.* **1991**, *16*, 378–381.
540. Sen, D.; Gilbert, W. Guanine quartet structures. *Methods Enzymol.* **1992**, *211*, 191–199.
541. Veselkov, A.G.; Malkov, V.A.; Frank-Kamenetskii, M.D.; Dobrynin, V.N. Triplex model of chromosome ends. *Nature* **1993**, *364*, 496.
542. Radding, C.M. Helical interactions in homologous pairing and strand exchange driven by RecA protein. *J. Biol. Chem.* **1991**, *266*, 5355–5358.
543. Stasiak, A. Three-stranded DNA structure: is this the secret of DNA homologous recognition? *Mol. Microbiol.* **1992**, *6*, 3267–3276.
544. Camerini-Otero, R.D.; Hsieh, P. Parallel DNA triplexes, homologous recombination, and other homology-dependent DNA interactions. *Cell* **1993**, *73*, 217–223.
545. Zhurkin, V.B.; Raghunathan, G.; Ulyanov, N.B.; Camerini-Otero, R.D.; Jernigan, R.L. A parallel DNA triplex as a model for the intermediate in homologous recombination. *J. Mol. Biol.* **1994**, *239*, 181–200.
546. Collier, D.A.; Griffin, J.A.; Wells, R.D. Non-B right-handed DNA conformations of homopurine-homopyrimidine sequences in the murine immunoglobulin C alpha switch region. *J. Biol. Chem.* **1988**, *263*, 7397–7405.
547. Kato, M. Polypyrimidine/polypurine sequence in plasmid DNA enhances formation of dimer molecules in *Escherichia coli*. *Mol. Biol. Rep.* **1993**, *18*, 183–187.
548. Rooney, S.M.; Moore, P.D. Antiparallel, intramolecular triplex DNA stimulates homologous recombination in human cells. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 2141–2144.
549. Kohwi, Y.; Panchenko, Y. Transcription-dependent recombination induced by triple-helix formation. *Genes Dev.* **1993**, *7*, 1766–1778.

550. Morgan, A.R.; Wells, R.D. Specificity of the three-stranded complex formation between double-stranded DNA and single-stranded RNA containing repeating nucleotide sequences. *J. Mol. Biol.* **1968**, *37*, 63–80.
551. Maher, L.J.I.; Dervan, P.B.; Wold, B. Analysis of promoter-specific repression by triple-helical DNA complexes in a eukaryotic cell-free transcription system. *Biochemistry* **1992**, *31*, 70–81.
552. Grigoriev, M.; Praseuth, D.; Guyesse, A.L.; Robin, P.; Thuong, N.T.; Helene, C.; et al. Inhibition of gene expression by triple helix-directed DNA cross-linking at specific genes. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 3501–3505.
553. Lu, G.; Ferl, R.J. Site-specific oligodeoxynucleotide binding to maize Adh1 gene promoter represses Adh1-GUS gene expression *in vivo*. *Plant Mol. Biol.* **1992**, *19*, 715–723.
554. Orson, F.M.; Thomas, D.W.; McShan, W.M.; Kessler, D.J.; Hogan, M.E. Oligonucleotide inhibition of IL2R α mRNA transcription by promoter region collinear triplex formation in lymphocytes. *Nucleic Acids Res.* **1991**, *19*, 3435–3441.
555. Postel, E.H.; Flint, S.J.; Kessler, D.J.; Hogan, M.E. Evidence that a triplex-forming oligonucleotide binds to the c-myc promotor in HeLa cells, thereby reducing c-myc mRNA levels. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 8227–8231.
556. Ing, N.H.; Beekman, J.M.; Kessler, D.J.; Murphy, M.; Jayaraman, K.; Zendegui, J.G.; et al. *In vivo* transcription of a progesterone-responsive gene is specifically inhibited by a triplex-forming oligonucleotide. *Nucleic Acids Res.* **1993**, *21*, 2789–2796.
557. Ojwang, J.; Elbaggari, A.; Marshall, H.B.; Jayaraman, K.; McGrath, M.S.; Rando, R.F. Inhibition of human immunodeficiency virus type 1 activity *in vitro* by oligonucleotides composed entirely of guanosine and thymidine. *J. Acquired Immune Def. Synd.* **1994**, *7*, 560–570.
558. Roy, C. Triple-helix formation interferes with the transcription and hinged DNA structure of the interferon-inducible 6-16 gene promoter. *Eur. J. Biochem.* **1994**, *220*, 493–503.
559. Duval-Valentin, G.; Thuong, N.T.; Helene, C. Specific inhibition of transcription by triple helix-forming oligonucleotides. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 504–508.
560. McShan, W.M.; Rossen, R.D.; Laughter, A.H.; Trial, J.; Kessler, D.J.; Zendegui, J.G.; et al. Inhibition of transcription of HIV-1 in infected human cells by oligodeoxynucleotides designed to form DNA triple helices. *J. Biol. Chem.* **1992**, *267*, 5712–5721.
561. Wang, Z.Y.; Lin, X.H.; Nobuyoshi, M.; Qui, Q.Q.; Deuel, T.F. Binding of single-stranded oligonucleotides to a non-B-form DNA structure results in loss of promoter activity of the platelet-derived growth factor A-chain gene. *J. Biol. Chem.* **1992**, *267*, 13669–13674.
562. Tu, G.C.; Cao, Q.N.; Israel, Y. Inhibition of gene expression by triple helix formation in hepatoma cells. *J. Biol. Chem.* **1995**, *270*, 28402–28407.
563. Kovacs, A.; Kandala, J.C.; Weber, K.T.; Guntaka, R.V. Triple helix-forming oligonucleotide corresponding to the polypyrimidine sequence in the rat $\alpha 1(I)$ collagen promoter specifically inhibits factor binding and transcription. *J. Biol. Chem.* **1996**, *271*, 1805–1812.
564. Kochetkova, M.; Shannon, M.F. DNA triplex formation selectively inhibits granulocyte-macrophage colony-stimulating factor gene expression in human T cells. *J. Biol. Chem.* **1996**, *271*, 14438–14444.
565. Xodo, L.E.; Alunni-Fabbroni, M.; Manzini, G.; Quadrifoglio, F. Pyrimidine phosphorothioate oligonucleotides form triple-stranded helices and promote transcription inhibition. *Nucleic Acids Res.* **1994**, *22*, 3322–3330.
566. Akhtar, S.; Kole, R.; Juliano, R.L. Stability of antisense oligonucleotide analogs in cellular extracts and sera. *Life Sci.* **1991**, *49*, 1793–1801.
567. Crooke, S.T. The future of sequence-specific transcriptional inhibition. *Cancer Invest.* **1996**, *14*, 89–90.
568. Lacroix, L.; Mergny, J.L.; Leroy, J.L.; Helene, C. Inability of RNA to form the i-motif: implications for triplex formation. *Biochemistry* **1996**, *35*, 8715–8722.
569. Olivas, W.M.; Maher, L.J., III. Competitive triplex/quadruplex equilibria involving guanine-rich oligonucleotides. *Biochemistry* **1995**, *34*, 278–284.

570. Fedoseyeva, E.V.; Li, Y.; Huey, B.; Tam, S.; Hunt, C.A.; Benichou, G.; et al. Inhibition of interferon-gamma-mediated immune functions by oligonucleotides. Suppression of human T cell proliferation by downregulation of IFN-gamma-induced ICAM-1 and Fc-receptor on accessory cells. *Transplantation* **1994**, *7*, 606–612.
571. Ramanathan, M.; Lantz, M.; MacGregor, R.D.; Garovoy, M.R.; Hunt, C.A. Characterization of the oligodeoxynucleotide-mediated inhibition of interferon-gamma-induced major histocompatibility complex Class I and intercellular adhesion molecule-1. *J. Biol. Chem.* **1994**, *269*, 24564–24574.
572. Gewirtz, A.M.; Stein, C.A.; Glazer, P.M. Facilitating oligonucleotide delivery: helping antisense deliver on its promise. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 3161–3163.
573. Leonetti, J.P.; Machy, P.; Degols, G.; Lebleu, B.; Leserman, L. Antibody-targeted liposomes containing oligodeoxyribonucleotides complementary to viral RNA selectively inhibit viral replication. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 2448–2451.
574. Bennett, C.F.; Chiang, M.Y.; Chan, H.; Shoemaker, J.E.; Mirabelli, C.K. Cationic lipids enhance cellular uptake and activity of phosphorothioate antisense oligonucleotides. *Mol. Pharmacol.* **1992**, *41*, 1023–1033.
575. Lewis, J.G.; Lin, K.Y.; Kothavale, A.; Flanagan, W.M.; Matteucci, M.D.; DePrince, R.B.; et al. A serum-resistant cytofectin for cellular delivery of antisense oligodeoxynucleotides and plasmid DNA. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 3176–3181.
576. Loke, S.L.; Stein, C.A.; Zhang, X.A.; Mori, K.; Nakanishi, M.; Subasinghe, C.; et al. Characterization of oligonucleotide transport into living cells. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 3874–3878.
577. Clarenc, J.P.; Degols, G.; Leonetti, J.P.; Michaud, P.; Lebleu, B. Delivery of antisense oligonucleotides by poly(L-lysine) conjugation and liposome encapsulation. *Anticancer Drug Des.* **1993**, *8*, 81–94.
578. Hentschel, C.C. Homopolymer sequences in the spacer of a sea urchin histone gene repeat are sensitive to S1 nuclease. *Nature* **1982**, *295*, 714–716.
579. Mace, H.A.F.; Pelham, H.R.B.; Travers, A.A. Association of an S1 nuclease-sensitive structure with short direct repeats 5' of *Drosophila* heat shock genes. *Nature* **1983**, *304*, 555–557.
580. McKeon, C.; Schmidt, A.; de Crombrugghe, B. A sequence conserved in both the chicken and mouse $\alpha 2(I)$ collagen promoter contains sites sensitive to S1 nucleases. *J. Biol. Chem.* **1984**, *259*, 6636–6640.
581. Pearson, C.E.; Sinden, R.R. Alternative DNA structures within the trinucleotide repeats of the myotonic dystrophy and fragile X locus. *Biochemistry* **1996**, *35*, 5041–5053.
582. Streisinger, G.; Okada, Y.; Emrich, J.; Newton, J.; Tsugita, A.; Terzaghi, E.; et al. Frameshift mutations and the genetic code. *Cold Spring Harb. Symp. Quant. Biol.* **1966**, *31*, 77–84.
583. Sheflin, L.G.; Kowalski, D. Mung bean nuclease cleavage of dA + dT-rich sequence or an inverted repeat sequence in supercoiled PM2 DNA depends on ionic environment. *Nucleic Acids Res.* **1984**, *12*, 7087–7104.
584. van de Sande, J.H.; Ramsing, N.B.; Germann, M.W.; Elhorst, W.; Kalisch, B.W.; von Kitzing, E.; et al. Parallel stranded DNA. *Science* **1988**, *241*, 551–557.
585. Jovin, T.M.; Rippe, K.; Ramsing, N.B.; Klement, R.; Elhorst, W.; Vojtiskova, M. Parallel stranded DNA. In: *Structure and Methods* (Sarma, M.H.; Sarma, R.H.; eds.). Adenine Press, Albany, NY, 1990, pp. 155–174.
586. Sen, D.; Gilbert, W. Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. *Nature* **1988**, *334*, 364–366.
587. Hardin, C.C.; Henderson, E.; Watson, T.; Prosser, J.K. Monovalent cation induced structural transitions in telomeric DNAs: G-DNA folding intermediates. *Biochemistry* **1991**, *30*, 4460–4472.
588. Williamson, J.R.; Raghuraman, M.K.; Cech, T.R. Monovalent cation-induced structure of telomeric DNA: the G-quartet model. *Cell* **1989**, *59*, 871–880.

589. Kang, C.K.; Zhang, X.; Ratliff, R.; Moyzis, R.; Rich, A. Crystal structure of four-stranded *Oxytricha* telomeric DNA. *Nature* **1992**, *356*, 126–131.
590. Smith, F.W.; Feigon, J. Quadruplex structure of *Oxtricha* telomeric oligonucleotides. *Nature* **1992**, *356*, 164–168.
591. Sen, D.; Gilbert, W. A sodium-potassium switch in the formation of four-stranded G4-DNA. *Nature* **1990**, *344*, 410–414.
592. Scaria, P.V.; Shire, S.J.; Shafer, R.H. Quadruplex structure of d(G₃T₄G₃) stabilized by K⁺ or Na⁺ is an asymmetric hairpin dimer. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 10336–10340.
593. Glover, J.N.; Pulleyblank, D.E. Protonated polypurine/polypyrimidine DNA tracts that appear to lack the single-stranded pyrimidine loop predicted by the “H” model. *J. Mol. Biol.* **1990**, *215*, 653–663.
594. Shimizu, M.; Hanvey, J.C.; Wells, R.D. Multiple non-B-DNA conformations of polypurine-polypyrimidine sequences in plasmids. *Biochemistry* **1990**, *29*, 4704–4713.
595. Johnson, D.; Morgan, A.R. Unique structures formed by pyrimidine-purine DNAs which may be four-stranded. *Proc. Natl. Acad. Sci. USA* **1978**, *75*, 1637–1641.
596. Ramsing, N.B.; Jovin, T.M. Parallel stranded duplex DNA. *Nucleic Acids Res.* **1988**, *16*, 6659–676.