1

DNA Structure: Alphabet Soup for the Cellular Soul

P. Shing Ho and Megan Carter
Department of Biochemistry and Molecular Biology,
Colorado State University
USA

1. Introduction

The story of DNA structure is as varied as it is interesting, the most famous tale being the "discovery" of B-DNA by Watson and Crick. For many biologists, this simple, but elegant structure is all that is needed for a basic, albeit superficial understanding of cellular genetics. A deeper appreciation for how DNA functions comes from the recognition that this is a highly malleable molecule, providing the cell with a plethora of conformations to exploit in replication and transcription. Some of these conformations can give rise to mistakes, while others help to repair those mistakes in the genetic code. In this chapter, we dive into the cellular pot and find a literal alphabet soup of DNA structures. We start our journey by presenting the fundamental principles that serve as the vocabulary to analyze and describe the features of nucleic acid structures. We will explore the conformational variations that lead from double-helices to complexes composed of three or four strands, then consider how conformations interconvert through various intermediates. Although B-DNA is the standard form in the cell, we suggest that this dance away from the norm is essential for cellular function, giving the cell life and, hence, its genetic soul.

Replication is the process by which the cell creates an exact copy of the genetic information encoded in DNA—it is thus intuitive that we would be interested in the actual structure of DNA as a molecule. One would think that, for replication, we need only be concerned with the DNA duplex at the beginning, the single-stranded intermediate state, and the final duplex, since these structures generally tell us how the information is stored and read, and what the resulting product is. However, it is becoming clear that although the general structure of DNA is important in the overall mechanism of replication (Watson & Crick, 1953a), the conformational details are important for understanding how proteins recognize their cognate DNA sequence, and how mutations may be introduced and are repaired. Thus, we must explore and dissect the details in terms of variations that define the particular sequence dependent shape of DNA.

We will not attempt the impossible task of covering every aspect of DNA structure, only those that may be relevant to replication. Also, as crystallographers, we will have a bias towards studies derived from X-ray diffraction and other physical methods, although we will always attempt to relate these back to the biology of replication. In the process, we will explore the details of DNA structure that help elucidate structural principles that contribute to our understanding of the mechanism and fidelity of the replicative process.

2. A brief history of DNA structure

DNA structure has had over 55 years of history and, in that time, has undergone periods of discovery that have pushed the field forward in spurts. The evidence that DNA is the genetic molecule in the cell came from the studies of Avery, MacLeod, and McCarty (Avery *et al.*, 1944), and confirmed by Hershey and Chase (Hershey and Chase, 1952). The seminal experiments of Meselsen and Stahl (Meselson and Stahl, 1958) using heavy atom labeled DNA demonstrated that replication is semiconservative, with each newly replicated daughter strand being paired with one of the two parental strands. These classic studies from the 1940's and 1950's set the stage for a race to determine the molecular structure of DNA, a now familiar story that helps to bring perspective to the discussions in this chapter.

2.1 The race for the structure of DNA: X-ray fiber diffraction studies.

The key element in the race towards the structure of DNA was the availability of X-ray diffraction photographs of DNA fibers, the best of which came from the work of Franklin and Gosling in the lab of John Randall. It was clear at the time that DNA could adopt two different forms, an A-form under low humidity and a B-form at higher humidity. The A-DNA form gave the highest resolution data (Franklin and Gosling, 1953a), but, it was the lower resolution photograph of the "wetter" B-form (Franklin and Gosling, 1953b) (Fig. 1) that was more readily interpretable. From this photograph, DNA was clearly seen to be a helical structure (showing the characteristic "helical-X" in the diffraction pattern), with a repeat of 10 units (reflected in the pattern converging after 10 layer lines), and with a distance between repeating units of 3.4 Å (from the *d*-spacing of 10th layer line). What was not evident was the number of strands in the helix (indeed, Linus Pauling had initially proposed a three-stranded structure (Pauling and Corey, 1953)), whether it is left- or right-handed, and how the information is read and properly replicated. The interpretation of this data by Watson and Crick (Watson and Crick, 1953b) lead to the iconic right-handed, antiparallel, double-helical model of DNA that we all recognize.

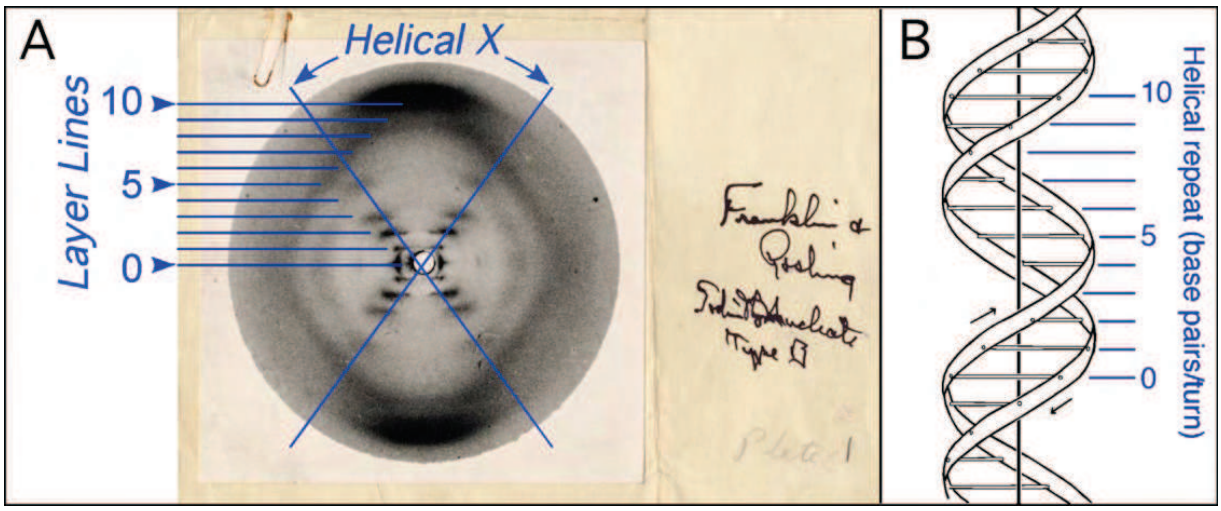

Fig. 1. Structure of B-DNA. A. Photograph 51 of B-DNA. X-ray diffraction photograph of a DNA fiber at high humidity (Franklin and Gosling, 1953b). Interpretation of the helical-X and layer lines added in blue. B. Watson-Crick model of B-DNA, adopted from (Watson and Crick, 1953b), with the helical repeat associated with the layer lines labeled.

Often missing from this story is that the Watson-Crick model depended not only on the large amount of biochemical and X-ray diffraction data being generated at the time, but also on a proper understanding of the chemical properties of DNA. One of the most important aspects of the Watson-Crick model was the proposal that guanines paired with cytosines and adenines with thymines. For this to occur, however, the nucleotide bases must be drawn in their proper tautomeric forms; however, up to that point, it was not clear, even to the organic chemists, what those forms should be. The initial assignment of guanine and thymine bases in their enol forms had lead to an early parallel model for DNA (Watson, 1968). It was not until the proper tautomers for the common nucleotides were assigned that the now familiar base pairs of G to C and A to T made sense, and, thus, provide a rationale for the well understood Chargaff rules for the complementary composition of nucleotides in the DNA of higher organisms (Chargaff, 1950) and a mechanism by which exact copies of the sequence information along a strand of DNA could result in an exact copy of a duplex through semiconservative replication (Watson and Crick, 1953a).

2.2 The single-crystal structures of DNA oligonucleotides

At this point, it should be stressed that Watson and Crick did not "discover" or "solve" the structure of DNA, but had presented a plausible and, basically, correct model that made important predictions that, in the end, led to the birth of modern molecular biology. However, several decades will pass before high resolution single crystals structures of synthetic DNAs emerge to support the essential elements of this model. For example, it was not immediately obvious that the Watson-Crick scheme, particularly for A=T base pairs, was correct—at the time the single-crystal structures of adenine bases paired with thymine or uracil bases showed geometries of Hoogsteen-type base pairs (this will be defined in Section 3). It was not until the crystal structure of the RNA dinucleotide phosphate ApU was determined to a remarkable 0.89 Å resolution (in crystallography, lower numbers refer to higher resolution) by Alexander Rich's group (Rosenberg *et al.*, 1973) that the Watson-Crick form of the A=U (and, thus, the analogous A=T) base pairs were confirmed. The concurrent structure of GpC also confirmed the Watson-Crick form of the G•C base pair (Day *et al.*, 1973) and, together, demonstrated for the first time that nucleotide double-helices (in this case, RNA dinucleotides) were antiparallel and had a right-handed twist.

In the late 1970's, it became possible to chemically synthesize "long" stretches of a defined DNA sequence for crystallographic studies. In 1979, Rich's group (Wang *et al.*, 1979) determined the single crystal structure of the DNA sequence CGCGCG (we write only one strand and drop the "p" for the phosphates for the sake of efficiency, even for double-helical structures). This structure showed DNA to be an antiparallel double-helix with Watson-Crick type base pairs, consistent with the 1953 model. However, it came with a new twist—this double-helix was left-handed and was called Z-DNA (for the zig-zagged backbone). It was not until 1981, with the single-crystal structure of the sequence CGCGTATACGCG (known as the Drew-Dickerson dodecamer (Drew *et al.*, 1981)), that the Watson-Crick structure for B-DNA was finally "proven" to be correct.

So, what of the dehydrated A-DNA form that Franklin had worked so hard on and struggled with? Soon after the Watson and Crick model of B-DNA, Franklin and Gosling published the structure of the fiber A-DNA form (Franklin and Gosling, 1953a), with a large number of single-crystals of A-DNA being determined and published in the 1980's and 1990's (the "heydays" of DNA crystallography (Mirkin, 2008)). The A-form was subsequently shown to be the native form of RNA duplexes, while DNA/RNA hybrids (primers for replication initiation) can interchange between the A- and B-forms.

Although it is well accepted that the B-DNA form is the most prevalent form in solution and in the cell, there is now a myriad of single-crystal DNA structures, including those assembled as double-, triple-, quadruple-, and even hexa- and octa-stranded complexes. There are hairpins from single-strands, structures with overhangs, etc., and a plethora of forms seen in complexes with proteins. We will discuss some of these in greater detail in Section 4 along with their relevant cellular functions, focusing on replication and the associated processes. First, we must delve into the detailed vocabulary used to describe DNA structure and provide a common language for the remainder of the chapter.

3. A vocabulary lesson for DNA structure

As with any description of a biopolymer, we will start the discussion of DNA structure at the simplest unit (the nucleotide building block), then develop the concepts of structure with increasing size and complexity. In order to reach this stage of complexity, we must first define terms that will be used in discussing DNA structure at all levels.

3.1 General principles

Almost every student today knows that DNA is composed of four basic building blocks, each defined by the unique chemical structure of the aromatic base, and each base attached to a phosphodeoxyribose backbone. The four common deoxyribonucleotides are categorized as the purine (deoxyadenosine, dA, and deoxyguanosine, dG) or pyrimidine (deoxythymidine, dT, and deoxycytosine, dC) nucleotides. The atoms of sugars are distinguished from those of the bases by a "prime" added to the atom name, so that the sugar carbons are C1′, C2′, C3′, C4′, C5′ (Fig. 2), starting with the carbon at the glycosidic bond that attaches the base to the sugar, and so forth around the ring. The deoxynucleotides of DNA lack a O2′ oxygen, which distinguishes them from ribonucleotides (RNA). For simplicity, we will simply assume the deoxyform and drop the "deoxy" and "d" prefixes from this point on (Hendrickson *et al.*, 1988).

3.2 What defines a stable DNA structure?

DNA in its functional form is not the isolated nucleotides, but a polymer built from the mononucleotides (G, C, A, T). A DNA polymer is constructed through condensation to form a phosphodiester linkage that bridges the O3′ and O5′ oxygens of sequential nucleotides (Fig. 2A). The primary structure, or sequence, of a DNA polymer strand is written in the direction that they are synthesized in the cell, starting at the free O5′ oxygen (5′-end) and progresses to the free O3′–end. Two complementary strands are brought together in a sequence specific manner to form an antiparallel double-strand, aligning one strand in the 5′ to 3′ direction and the complement 3′ to 5′. Nearly all functional secondary structures of DNA are multi-stranded, most commonly double-stranded. As the sequence of one strand dictates that of its complement, double-stranded DNA is often considered as a single biological molecule, even though the strands are not covalently linked.

3.2.1 Base pairing

Unlike proteins and RNA, the functional forms of DNA are typically complexes comprised of two or more strands, which are stabilized by base pairing, base stacking, and solvent interactions. Of these, base pairing is best understood for its important role in specifying the

sequence of newly synthesized DNA during replication and in general sequence recognition, but is perhaps the most misunderstood for its contribution to DNA stability.
The most commonly recognized form of DNA, B-DNA, is the double-stranded duplex stabilized by Watson-Crick base pairing (Fig. 2B). In standard Watson-Crick G•C and A•T base pairs, hydrogen bonds are formed between the respective donor and acceptor functional groups along what is called the "Watson-Crick" edges of the bases. The geometries of both purine-pyrimidine base pairs are similar in the relative positions of their bases and, consequently, the width of the resulting major and minor groove—the similarity in the geometries of correctly paired bases contributes to the fidelity of the replication polymerases [(Kool, 2001)]. The G•C base pair, however, is stabilized by three hydrogen bonds as opposed to the two that stabilize A•T base pairs; thus, G•C rich sequences tend to have higher stabilization energies and melting temperatures. With only two hydrogen bonds, A•T base pairs offer less resistance to deformations, including twisting of the individual bases from a common plane (called propeller twist, see below). Although the standard Watson-Crick base paired duplex DNA is most universally recognized, it is clear that DNA structures with non-standard pairing of bases are more prevalent and biologically significant than previously thought [(Neidle, 1999)].

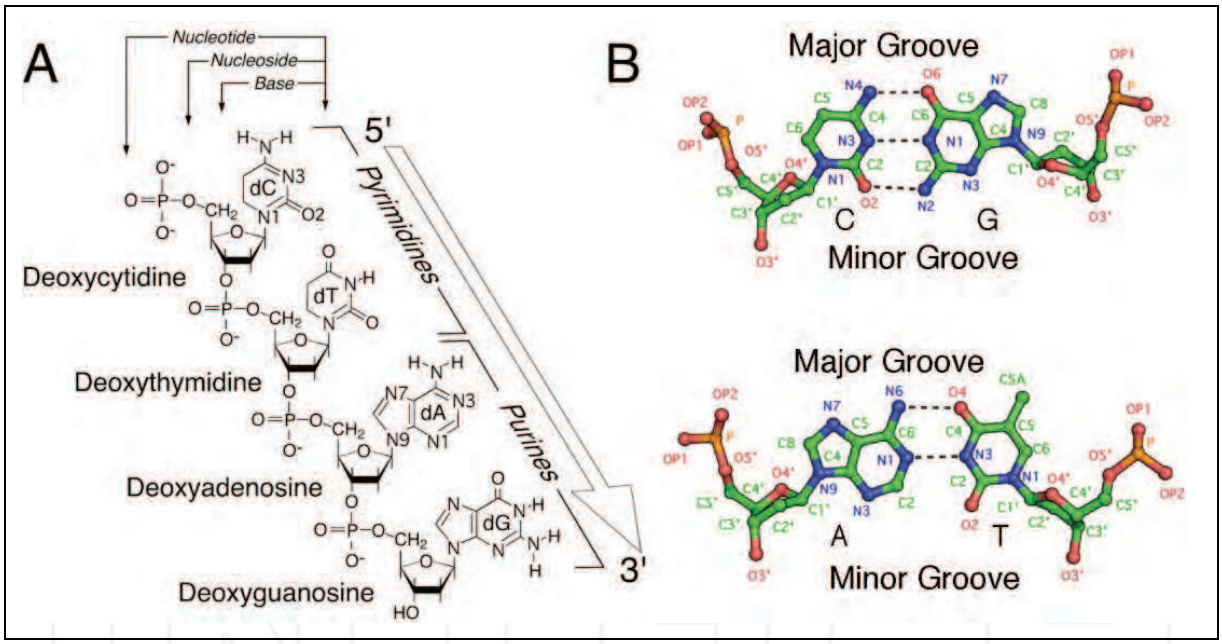

Fig. 2. Components of DNA. A. The four common deoxyribonucleotides are connected by phosphodiester bonds to form a single-strand, 5' to 3'. B. Watson and Crick C•G and A•T base pairs with the major and minor grooves labeled.

Non-standard base pairs play critical roles in the varied structures observed in DNA and RNA. Wobble, mismatched, and reverse base pairs still use the Watson-Crick edges for hydrogen bonding. Reverse Watson-Crick base pairs are found in parallel duplexes, but are not immediately relevant to DNA replication. Wobble base pairing (Fig. 3A) is seen in mismatches between G•T and G•U base pairs incorporated into DNA and DNA:RNA complexes and play essential roles in the fidelity of DNA replication and transcription. Such mismatches can lead to genome mutations if not accurately detected and corrected by the

proof reading activity of DNA polymerase during replication, or post-replicative repair systems. Studies suggest that G•T and A^+•C are the most frequent mismatches that cause point mutations in cells (Neidle, 1999). The energies of hydrogen bonding in proper and mismatched bases, relative to base stacking and steric effects, however, appear to have little influence on polymerase fidelity (Kool, 2001).

Hoogsteen base pairs take advantage of the Hoogsteen edge of a purine base, which is orthogonal to and, thus, can be accessed without disrupting the Watson-Crick base pairing edge (Fig. 3B). Consequently, Hoogsteen interactions allow the assembly of multi-stranded DNA complexes, including triplet helixes and G-quadruplexes.

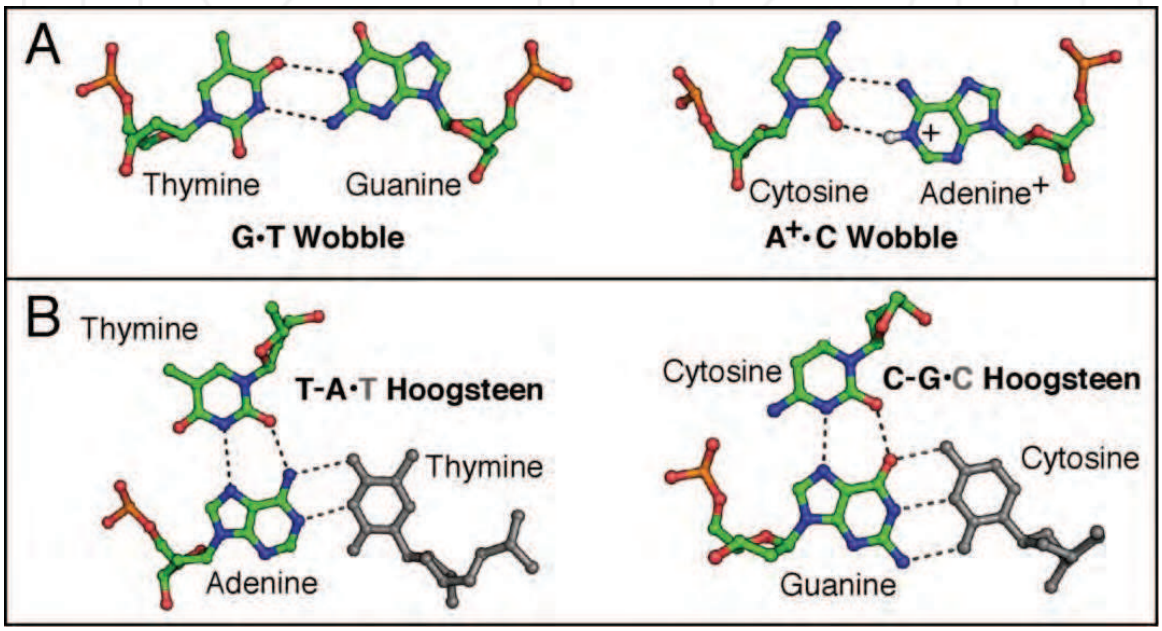

Fig. 3. Non-Watson Crick base pairs. A. G•T wobble and A^+•C wobble base pairs. B. Thymine Hoogsteen paired to A•T WC base pair, cytosine Hoogsteen paired to G•C WC base pair as observed in triplex strand formation.

3.2.2 Base stacking

Although not as intuitive, the stacking of bases into a column is as or more critical to the stability of multistranded DNAs (duplexes, triplexes, tetraplexes, etc) as base pairing. It is estimated that base stacking contributes as much as half of the total stabilizing free energy of a base pair in duplex DNA (Kool, 2001). Van der Waals interactions, electrostatic interactions, and solvent effects define the geometry and associated energies of stacked bases. Van der Waals forces drive bases to stack in a way that best complements their surface topologies. In addition, individual atoms carry permanent partial charges that contribute to either Coulombic attraction or repulsion between bases. This can be modeled as interactions between permanent dipoles, and it is this dipolar interaction, in conjunction with shape complementarities that helps to define the orientation of the stacked bases. The specific orientation of stacked base pairs contributes to the conformational stability of a DNA duplex. Likewise, deformations associated with specific base stacking geometries contribute to the mechanism of indirect sequence specific binding and recognition by proteins. Finally, since the nucleotide bases are aromatic and, therefore, primarily hydrophobic, stacking minimizes the solvent exposure of the base surfaces, thus, leading to

the familiar face-to-face stacking of bases and base pairs. It is not surprising, therefore, that DNA conformations that increase exposure of bases are stabilized by organic solvents.

3.2.3 The phosphodeoxyribose backbone

The functional form of DNA links nucleotides together by phosphodiester bonds to form a continuous DNA strand. Phosphodiesters are highly acidic (pKa' ~1.5); thus, at neutral pH, the phosphate group is a monoanion with a formal -1 charge distributed among all four oxygens, with the two non-ester oxygens (OP1, OP2) carrying about twice the charge as the ester bonded oxygens (O5', O3'). As a consequence, the DNA phosphoribose backbone is overall negative and provides an opposing force to the base pairing and stacking interactions that hold a DNA duplex together. Indeed, if the backbone were uncharged, it would be much more difficult to unzip or displace a DNA strand and, consequently, it would take more energy to unwind a duplex to allow replication to start and to proceed.

The overall charge of DNA in solution is not simply a sum of -1 for each nucleotide—the backbone charges are counterbalanced by positive cations that accumulate around the DNA. These counterions are simple ions (monovalent Na^+ and K^+, or divalent Mg^{+2} and Ca^{+2} being the most prevalent in a cell), but include cationic polyamines (spermine and spermidine), drugs (ethidium or *cis*-platin), or proteins (*e.g.,* the histone proteins of nuclesomes). In general, DNA in solution is less negatively charged than expected—as a polyelectrolyte, each phosphate of a DNA duplex carries an "effective" charge of approximately -0.6, or that ~40% of the charge is counterbalanced by simple cations (Manning, 1977). The remaining net charge, however, acts to destabilize the double-helix. Consequently, structures with closely spaced phosphates are stabilized by increased concentrations of counter cations.

When a protein, such as DNA polymerase, binds to DNA, it must competitively displace the counterions associated with the DNA backbone. For example, nucleosome formation, which helps compact DNA in eukaryotes, is primarily driven by nonspecific interactions of the positive histones with the negative DNA backbone. In order to replicate or transcribe the information of the DNA, the respective polymerase and all of its associated proteins must compete against these non-specific interactions. Thus, the negative charge of the backbone is a platform for sequence independent electrostatic interactions with proteins in the cell (Rohs, et al., 2009).

3.2.4 Solvent effects

As with any biological molecule, solvent interactions directly influence DNA structure and function. Base pairing and stacking are in part stabilized by the hydrophobic effect. We have already seen how solvent (considered to consist primarily of water and salts) induces base pairs to stack and defines the effective charge of the phosphoribose backbone. Even base pairing is affected by solvent interactions. In forming a base pair, the hydrogen bond donor and acceptor groups of each base must break hydrogen bonds with water molecules first. If the enthalpy of any single hydrogen bond from one base to another base is essentially the same as they are from the base to water, why then do bases pair and exclude water (at 55.5 M concentration)? The primary answer is that sequestering hydrogen-bonding groups from the competing interactions of water increase the hydrogen bonding potential (Klotz, 1962). One can see from this why base stacking is so important in stabilizing double-, triple, and other multistranded DNA forms that are assembled through hydrogen bonding.

Water, however, is not entirely excluded from, but plays an important role in the structure of DNA. Even in a fully base paired duplex, numerous hydrogen bond donor and acceptor

groups of the backbone and bases must be hydrated. There are classes of waters that can, in fact, be considered integral components of a DNA's structure. In a G•T wobble base pair, for example, the number of hydrogen bonds between the bases is reduced by one; however, bridging water molecules help to compensate for this loss (Ho *et al.*, 1985). Similarly, there are well-defined waters lining the minor groove of B-DNA duplexes (the so-called "spine of hydration") (Drew *et al.*, 1981) that exchange slowly with the bulk solvent (Liepinsh *et al.*, 1992) and, therefore, are considered to be integral parts of DNA. Thus, water promotes base stacking, which provides an environment for more stable hydrogen bonds within base pairs. Waters solvate the surfaces of the major groove and form well defined hydrogen bonded networks that bridge the two strands across the minor groove. In order to minimize the opposing repulsion between the phosphates of the DNA strands, cations help to mitigate the negative charges of the phosphoribose backbone (Hamelberg *et al.*, 2001). It is evident, therefore, just how important solvent really is for DNA structure and stability.

Finally, we must briefly discuss how solvent plays a role in DNA function. DNA is a hydrated molecule, until it is bound to a protein, at which point the DNA becomes dehydrated—*i.e.*, a protein must compete against water in order to bind to the DNA. The basic concept of direct read-out of DNA base pairs is a prime example of this. Direct read-out requires a protein to essentially stick its hydrogen bonding side-chain fingers into places where they would not normally belong, the major groove of a DNA duplex, for example. Both the proteins side chains and the DNA surface that they are trying to read would prefer to remain solvated; however, in order to form a strong complex with DNA, the protein must expel water from both surfaces and, as a result, the complex will become more stable than the sum of the individual parts. This, again, requires a balance between the stability of hydrogen bonds, the resulting decrease in conformational entropy of the protein side chains, and an increase in entropy of the water molecules as they return to the bulk solvent.

3.3 Conformations of the deoxyribose sugar

In addition to charge effects, the phosphoribose backbone helps to define the conformation of DNA *via* the conformation of the deoxyribose sugar. The detailed conformation of any polymer is defined by the rotations about each freely rotating chemical bond (Fig. 4A). We can define three categories of bonds: those of the phosphodiester holding two nucleotides together, those within the five-membered ring of the deoxyribose sugar, and the bond holding the nucleotide base to the sugar. The angles around the bonds that hold two nucleotides together start at the oxygen that links phosphate to the C5′-carbon of the ribose ring. Rotation about the P-O5′ bond is the α-torsion angle, which is followed by the β-angle for the O5′-C5′ bond, and so forth until we get to the ζ-angle that links the O3′-oxygen to the phosphate of the next nucleotide. These bonds adopt angles that help to minimize the repulsion of the negatively charged phosphates within and between DNA strands.

The bonds in the furanose ring are distinguished from those that flow linearly from one nucleotide to the next, and are designated as ν_1 for the C1′-C2′ bond, ν_2 for the C2′-C3′ bond, and so forth (Fig. 4A). The reader would recognize that the ν_3 angle within the ring coincides with the δ-angle along the chain. The ring is non-planar, and it is how particular atoms are placed either above or below a reference plane (the "sugar pucker") that facilitates formation of various conformational forms of DNA. The torsion angles are correlated to maintain reasonable bond lengths and angles within the ring, and are described by a single *pseudorotation angle* Ψ, which defines the sugar pucker (Saenger, 1984). Sugars with atoms puckered above the reference plane (on the same side as the base) are in an *endo*-form (C2′-

endo pucker has the C2′-carbon pointed up and towards the base), while a pucker that places an atom below this plane is in its *exo*-form (Fig. 5). The two general classes of sugar conformations commonly seen in DNA are the C2′-*endo* and C3′-*endo* puckers—the interconversion between these forms will be discussed in detail in section 5. The two conformations have profound effects on the overall DNA conformation in that they specify different phosphate-phosphate distances along each strand (~7 Å for C2′-*endo* and ~6 Å for C3′-*endo*). Thus, conformations constructed with C3′-*endo* sugars will require higher concentrations of salts to counter balance the shorter distance between the negatively charged phosphates.

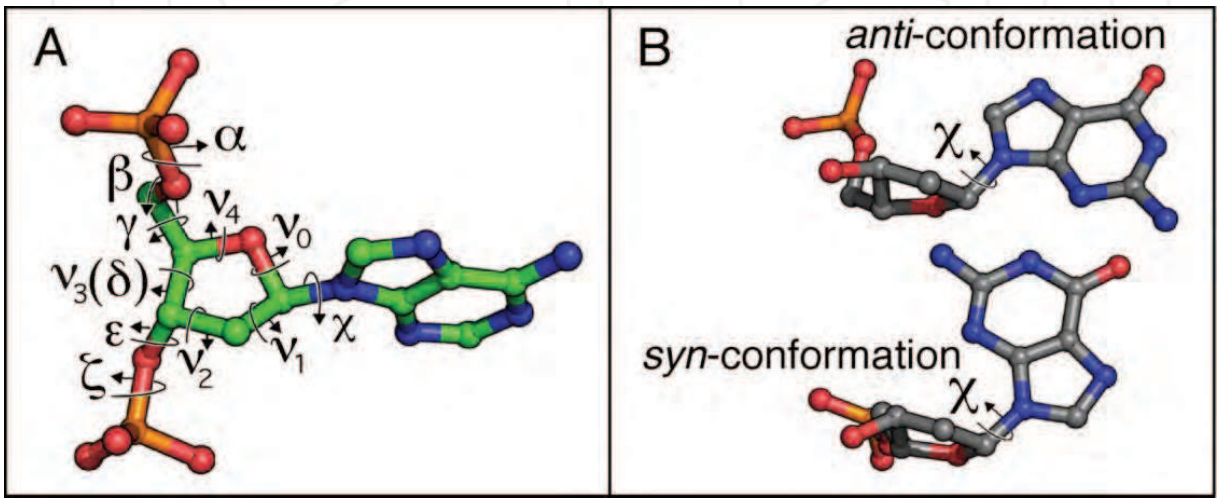

Fig. 4. Torsion angles of nucleic acids. A. Torsion angles along the backbone (α to ζ), within the sugar ring (ν_0 to ν_4), and the rotation of the nucleobase relative to the sugar. B. Rotation about the glycosidic bond defines χ-angles for the *anti*- and *syn*-conformations of the bases.

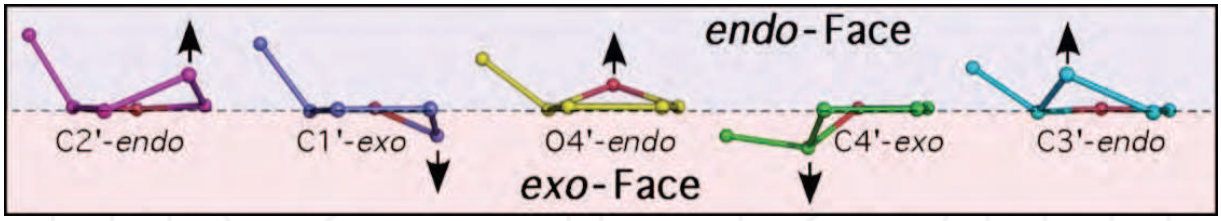

Fig. 5. Sugar pucker. Shown are the *endo* (above) and *exo* (below) faces of the 5-membered furanose sugar with the nucleotide base extended above the reference plane. Sugars are shown in order of transformation from C2′-*endo* to C3′-*endo*. Arrows indicate the atom that is puckered, and the direction of puckering.

The base of each nucleotide is attached *via* the glycosidic bond from the N1 nitrogen of pyrimidines or the N9 nitrogen of purines to the C1′-carbon of the deoxyribose sugar. The rotation about the glycosidic bond, the χ-angle, defines two general conformational classes: the *anti* conformation ($+90° \le \chi \le +180°$), with the base extended away from the sugar, and the *syn* conformation ($-90° \le \chi \le +90°$), with the base essentially lying on top of the sugar ring (Fig. 4B). The more compact *syn*-conformation is more susceptible to steric clashes than the extended *anti*-form. Although purine rings are generally larger, it has the smaller five-membered ring, as opposed to the six-membered ring of pyrimidines, attached directly to the sugar. Thus, purines will more readily adopt the compact *syn*-conformation than pyrimidines, because of reduced

steric collisions. Similarly, the *syn* conformation is less sterically hindered when the sugar is puckered as C3′-*endo* than C2′-*endo*. From this, we can now start to appreciate how the interplay between sugar puckers and χ-rotations can have profound effects on the structures of DNA and the sequence dependence for their formation.

3.4 Helical parameters

Now that we have assembled well-defined helical structures, how do we describe these structures? We can certainly do this in a very descriptive and qualitative manner, using the classical A- and B-forms as examples. For instance, we can characterize the standard B-form of DNA as a right-handed double-helix held together by Watson-Crick type base pairs that stack directly along a helical axis, resulting in two well defined grooves. However, this raises numerous questions, for example, at which point does a distortion to the Watson-Crick base pair become a wobble base pair, how far off the helix axis is allowed in this definition, and what if the helix axis is not straight? To address these and other questions, a set of quantitative measures called the "helical parameters" were developed to characterize the regular secondary structures of nucleic acids (both DNA and RNA) (Lavery, 1998).

The most commonly recognized parameters for DNA include the helical repeat (number of base pairs in one complete turn) and the helical rise (distance between nucleotides when measured along the helical axis). The repeat defines the angle relating each base pair along the helix axis (the helical twist = 360°/repeat), while the product of repeat and rise is the pitch (distance between one complete turn) of the DNA. These parameters restrict the geometries of the DNA. Indeed, if we consider only the closest physical approach between base pairs (the rise = 3.4 Å, as defined by the thickness of a base), the maximum phosphate-phosphate distance along a strand (measured at ~7.5 Å by single-molecule stretching (Allemand *et al.*, 1998)), and the effective diameter of a duplex (9.5 Å), we see that the largest twist angle between stacked base pairs is ~42°, resulting in a smallest theoretical repeat of 8.5 base pairs per turn. This would be the most tightly or over-wound form of a DNA double-helix. If the phosphate-to-phosphate distance is relaxed to ~7 Å (for a C2′-*endo* sugar pucker), the helical twist becomes ~36°, which translates to the ~10 bp/turn repeat of B-DNA. Finally, if the sugar adopts a C3′-*endo* conformation with a ~6Å phosphate-to-phosphate distance, the result is a structure with a helical twist of ~31° and a repeat of 11 – 12 base pairs, similar to that of A-DNA. We can see, therefore, how the sugar pucker defines the intrastrand phosphate-to-phosphate distance, base stacking defines the base-to-base distance, the base pairs define the radius of the DNA, and, finally, how all this comes together to define the way the DNA double-helix twists into a specific conformation. Of course, these are only very rough approximations of DNA structures—the detailed descriptions require a complete set of helical parameters in addition to the two described so far.

The helical parameters can be categorized into two general classes to describe the absolute and relative conformations in nucleic acids (Fig. 6); base-pair parameters (for single base pairs) and base step parameters (for adjacent base pairs). We note that these classes are not mutually exclusive, but are interrelated. Twist and rise are clearly base step parameters, since they describe the relative angle and distance between two adjacent stacked base pairs. The other base-step parameters that are generally considered relevant include slide, roll, tilt, and shift. It is easy to see that slide can effectively increase the diameter of a DNA duplex and, consequently affect the helical twist and repeat. A-DNA, for example, shows a large slide between base pairs, while B-DNAs have small slides, placing the base pairs essentially stacked on top of each other. Not surprisingly, therefore, A-DNA has a larger overall diameter and, in fact, appears to have a hole down the middle when viewed down its helical axis.

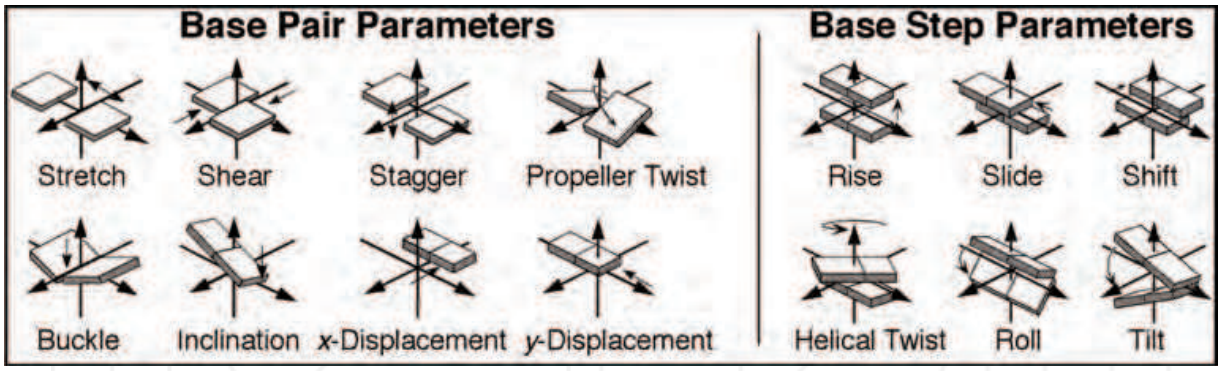

Fig. 6. Base Pair and Base Step Parameters. *Base Pair Parameters*: Translational and rotational relationships of bases within each base pair. *Base Step Parameters*: Translational and rotational relationsips between two stacked base pairs.

A conundrum in A-DNA is that it has a rise of ~2.5 Å, which would appear to violate the closest approach between stacked base pairs. In this case, the inclination associated with the roll and tilt of the base pairs, in conjunction with the helical twist result in a shortening of the vertical distance between base pairs along the helical axis, even though the stacking distance remains 3.4 Å. Indeed, A-like DNAs that have little or no roll and tilt have helical rises that are ~3.4 Å, as expected (Ng *et al.*, 2000; Vargason *et al.*, 2001).

Base pair parameters include those that relate the position or orientation of the base pair relative to the helical axis (inclination, *x*-displacement, and *y*-displacement), or the orientation and positions of the two bases in a pair (propeller twist, shear, stagger, stretch, buckle). It should be obvious that the inclination of a base pair will strongly influence the roll and tilt between base pairs, while slide defines the displacement perpendicular to the base pair (*x*) and along the base pair (*y*). Within the base pair itself, the large propeller twist seen in A•T base pairs has been attributed to the flexibility of two hydrogen bonds relative to three observed in G•C base pairs. At the extreme, this results in bifurcated hydrogen bonds, which are considered to be shared between adjacent A•T base pairs (Coll *et al.*, 1987).

Each of these base pair and base step parameters are defined relative to the helical axis that runs down the center of DNA. However, it should be recognized that defining this axis is not entirely straight forward, particularly if the DNA trajectory is bent or curved. There are two approaches to defining helical axes: the global axis and the local axis. The global axis is essentially the continuous curve that best runs down the center of all base pairs in a structure, while the local axis is the best line that defines the center of any two adjacent base pairs (local axes need not be continuous). Thus, helical parameters are analyzed in the context of global or local axes, and are not interchangeable and may be very different.

Two distinguishing features of double-helical DNAs are the grooves. The widths of the major and minor grooves are measured as the phosphate-to-phosphate distance across the two strands in a direction perpendicular to the trajectory of the strands. These groove widths provide an important means for proteins to interact with the base pairs of the DNA. The wide major groove of B-DNA allows direct read-out of the bases, while the narrow major groove of A-DNA does not—there is, however, an advantage to A-DNA having a wider minor groove, which we will discuss in the next section. It should be immediately obvious from the earlier discussion that the base pair and base step parameters described above conspire to define the groove widths for each form of DNA.

Finally, we can see how a parameter such as twist has such a strong effect on the overall behavior of genomic DNAs. DNA when confined in the cell or the cell's nucleus must be

packaged into a compacted supercoiled form and, in the process, this induces stress that will perturb its secondary structure. For simplicity a set of terms have been defined for supercoiled DNA in the context of closed-circular double-stranded DNA such as those found in plasmids, bacterial chromosomes, and viral genomes. These terms can also be applied to linear eukaryotic DNAs that are spatially anchored and stressed through protein binding, DNA unwinding, and DNA compaction. In double-stranded DNA, the number of times the strands wrap around each other along the helical axis is defined as the twist (*Tw*), with positive *Tw* associated with right-handed and negative *Tw* for left-handed duplexes, and unwound duplexes (*e.g.,* melted domains) as *Tw* = 0. In closed-circular DNA, the ends are joined and not free to turn in accommodating a change in *Tw*; therefore, a change in twist has additional global effects (Fig. 7), resulting in supercoiling, or writhing (*Wr*), of the double-helix as it wraps around itself.

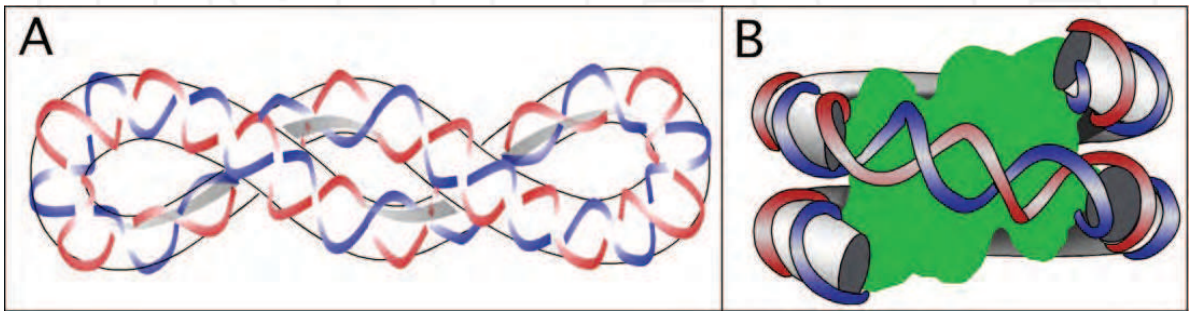

Fig. 7. Supercoiled DNA. A. Two negative supercoils are manifest as right-handed cross-overs in closed circular DNA. B. Similarly the DNA found in the nucleosome structure is wrapped ~twice around the histone core proteins (green) to form two negative supercoils (adapted from (Luger, *et al.,* 1997)).

Together, the twist and writhe define the topological properties of DNA. In truly closed-circular DNA that is unconstrained, twist and writhe are entirely correlated through the linking number (*Lk*) according to the equation *Lk* = *Tw* + *Wr*. Thus, if we unwind (reduce *Tw*) in closed circular DNA, the resulting strain must be relieved by increasing *Wr* (supercoiling). The only way to change *Lk* is by breaking the bonds of the backbone of one or both of the DNA strands, a process carried out by topoisomerases in the cells. How does all of this play out during replication? Consider the closed circular genome of a bacterium, or a domain of a eukaryotic genome that is locally constrained by nucleosomes and/or matrix attachment regions (MARs). As a DNA helicase plows through the DNA, it will locally unwind and melt the duplex (reduce *Tw*) for synthesis of the daughter strand. In doing so, the DNA in front of the polymerase will be positively supercoiled, while negative supercoils accumulate in its wake, both energetically unfavorable conditions. To relieve the strain, topoisomerases must relax the supercoils both in front of and behind the replisome.

4. The alphabet soup of DNA structures

DNA is highly polymorphic and, at least at the level of the helical structures, more variable than either proteins or RNA. The various forms of DNA have traditionally been named using the letters of the English alphabet and, from a survey of the literature, it was found that all but four letters have been assigned to at least one unique structural form (Ghosh and Bansal, 2003). We will, in this section, briefly describe a subset of DNA conformations that

have been structurally characterized (Fig. 8 and 9) and the sequence propensities of these structures, starting with B-DNA and working our way through the variations on the double-helix and various multi-stranded conformations. Along the way, we will discuss their potential biological functions, particularly in DNA replication, as appropriate.

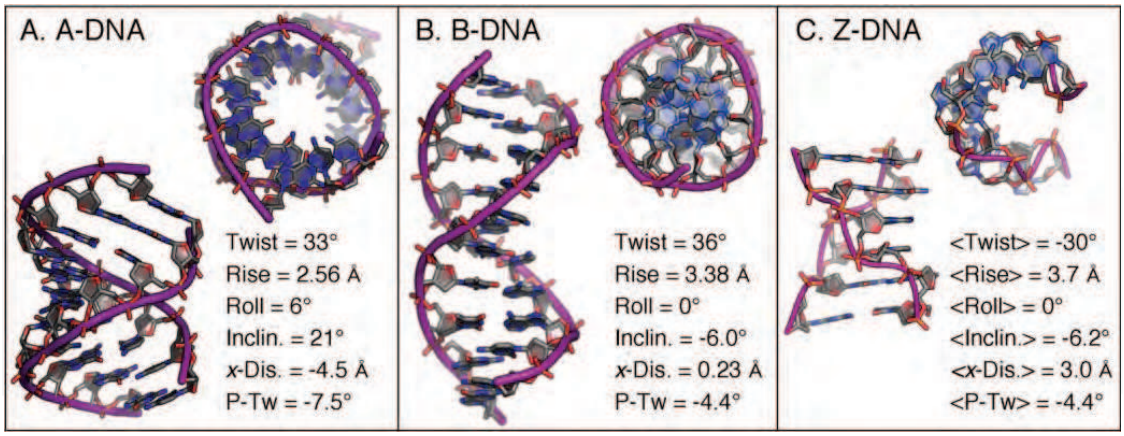

Fig. 8. Representatiive double-helical structures of DNA. Structures of A-DNA (Hays *et al.*, 2005), B-DNA (Privé *et al.*, 1991), and Z-DNA (Wang *et al.*, 1979). Abbreviations: Incl. = inclination, *x*-Disp. = *x*-displacment, P-Tw = propeller twist. For Z-DNA, the helical parameters are given as averages of the alternating dinucleotide steps.

4.1 B-DNA: The standard form

B-form DNA is the most recognized and common structural form of DNA in the cell, being considered the conformation adopted by nearly all sequences within a genome. Interestingly, while B-DNA has a distinguishing set of structural properties, it is now understood to be highly variable and malleable. B-DNA is a right-handed, antiparallel double-helix in which the Watson-Crick base pairs are stacked directly along and perpendicular to the helical axis, giving rise to major and minor grooves that are similar in depth. The bases are all in the *anti*-conformation with a majority of deoxyribose sugars in the C2′-*endo* form, although the sugar puckers are more variable than in many other conformations (Dickerson, 1999). The highly accessible major groove allows for direct readout of the polynucleotide sequence by proteins through patterns of hydrogen bond donors and acceptors that are complementary between the amino-acid side chains and each individual base pair. The more narrow minor groove, on the other hand, is characterized by a series of strongly coordinated waters and ions.

Although these properties are general for B-DNA, the structure is highly variable from one sequence to the next and for the same sequence under different conditions. The concept of *sequence-based differential deformability* recognizes that the B-form of a single sequence can adopt multiple conformations in response to the environment which can affect protein recognition. Therefore, the effect of sequence is important not in terms of any one structure, but instead in its malleability—the ability of that sequence to be deformed and molded as necessary for a particular function. For example, A•T base pairs and long stretches of A/T sequences (A-tract DNAs) seem to deviate significantly from the standard B-structure, showing larger propeller twists, along with narrower and more variable minor groove widths. Narrow minor grooves are shown to have preferential binding by arginine side chains of multiple DNA-binding protein families (Rohs *et al.*, 2009), and represent a specific

example of protein recognition based on sequence perturbations to standard B-DNA. A-tract DNA sequences are also associated with large rolls and tilts of their base steps, resulting in rigid bending of the B-DNA duplex (Neidle, 1999). An extreme example of these perturbations is seen with the structure induced in gene promoter sequences by the TATA-binding protein in transcription (called TATA-DNA), which shows a significant tilt and roll of the base pairs, unwinding of the duplex, and widening of the minor groove in a manner similar to that seen with A-DNA (Burley, 1996).

Variations of the B-form have been primarily elucidated by detailed structural studies, particularly X-ray diffraction and NMR, on short oligonucleotides. The question that is often raised is whether these short lengths of DNA may in fact not be relevant (and, in the case of crystals, be otherwise distorted (Dickerson *et al.*, 1994)) relative to sequences embedded in a genomic context. Studies by Tullius' group using hydroxyl-radical foot printing (Greenbaum *et al.*, 2007), have shown significant sequence dependent variation in the solvent accessibility and, thus, the helical structure of protein-free genomic DNA. These structural variations at the genomic level are highly correlated with variations in helical parameters measured in DNA crystal structures (unpublished results) derived from a self-consistent data set (Hays *et al.*, 2005). In conclusion, there is growing recognition that even B-DNA is a highly variable structural form of the DNA double-helix, and that sequence dependent structural variations play a critical role in protein recognition and binding.

4.2 A-DNA: Underwinding for replication fidelity

A-form DNA is also a right-handed antiparallel helical duplex, but is characterized as an underwound structure that is more compact along the helix axis and broader overall across the helix relative to B-DNA. The nucleotide bases, all *anti*, are shifted by large x-displacements towards the minor groove, creating a shallow, wide minor groove and a channel associated with a deep, narrow major groove. The deoxyribose sugars are consistently C3'-*endo*, which minimizes the potential steric clashes as the sugar is pushed towards the phosphate to accommodate the sliding of the base (Dickerson, 1999).

A-DNA is involved in insuring the fidelity of DNA replication. An analysis of the structure of the *Bacillus* DNA polymerase in complex with duplex DNA showed a conformational switch from the B- to underwound A-form starting at the site of nucleotide incorporation and extending to four bases upstream (Kiefer *et al.*, 1998). Why is A-DNA induced by the polymerase? There are several perspectives on this answer, from an evolutionary view (the emergence of DNA polymerase from the primoidial RNA world where RNA polymerase reigned) to a functional view. We will discuss the latter in slightly greater detail. The direct read-out mechanism involves sticking amino acid side-chains into the DNA's major groove to read the unique pattern of hydrogen bonding donors and acceptors that specify a particular sequence. One would think that this would be a fairly straight forward way for a polymerase to insure the fidelity of the newly synthesized daughter strand and, thus would want the double-helix to adopt the standard B-form with its wide and accessible major groove. However, DNA polymerases are not sequence specific (*i.e.,* they will synthesize from any template sequence), so the enzyme must distinguish a proper Watson-Crick base pair from various mismatches without knowing what the base pair should be. The characteristic feature of mismatched bases (as in a wobble) is that the structure of the minor groove becomes perturbed (Kool, 2001); thus, by inducing the A-form, the polymerase exploits the structural features of the highly accessible minor-groove to insure that the correct base has been added relative to the template sequence.

4.3 Z-DNA: The left-handed duplex

Z-form DNA is noteworthy as the only characterized left-handed form of the double-helix. The zig-zagged backbone, its namesake, results from the alternation between *syn*- and *anti*-conformations, and the respective C3′-*endo* and C2′-*endo* sugar puckers. This alternating conformation imposes a sequence preference for alternating purine-pyrimidines, since purines adopt the *syn*-conformation more readily than do pyrimidines. Thus, the repeating unit is the dinucleotide rather than a single base pair, as in B-DNA. The major groove in Z-DNA is not so much a groove but more a convex outer surface, while the minor groove becomes a deep, narrow and largely inaccessible crevice (Wang *et al.*, 1979).

The biological function of Z-DNA has been widely debated and underappreciated; however, several cellular functions for the Z-form are now supported by experimental evidence (Rich and Zhang, 2003). Z-DNA was initially characterized as a structure induced by high salt conditions (3 M NaCl) (Pohl and Jovin, 1972), leading many to wonder whether it could exist in a cell. Subsequently, it has been shown that cytosine methylation, and other cations such as spermine and spermidine at millimolar concentrations also stabilize Z-DNA (Rich and Zhang, 2003). Most importantly, as a left-handed structure, Z-DNA is the most underwound form of the double-helix and, consequently, serves as a sink for the torsional tension in negatively supercoiled DNA (Rich and Zhang, 2003). This expands the range of cellular situations that could support the formation, at least transiently, of Z-DNA. In one model, RNA polymerase, as it transcribes through a gene, would generate negative supercoils in its wake (Liu and Wang, 1987) and, on the process drive Z-DNA formation upstream of the transcribing gene. A detailed study of the promoter for human CSF-1 gene showed that up-regulation by the chromatin remodeling BAF protein involves a Z-DNA element (Liu *et al.*, 2001). The authors suggested that Z-DNA upstream of the nuclear factor-1 binding site helped to maintain the gene in its activated, nucleosome-free state (nucleosomes do not bind to the very rigid Z-DNA form (Ausio *et al.*, 1987)). In support of its potential role in the regulation of eukaryotic genes, we have found that Z-forming sequences accumulate near the transcription start site of genes in humans and other eukaryotes (Khuu *et al.*, 2007; Schroth *et al.*, 1992), and that ~80% of the genes in human chromosome 22 have at least one Z-DNA sequence in the vicinity of their transcription start sites (Champ *et al.*, 2004).

The discovery of protein domains having very high specificity for Z-DNA (Rich and Zhang, 2003), in some cases with nanomolar K_D's, have suggested additional functions that include, for example, RNA editing and gene transactivation. Z-DNA sequences have also been implicated in genomic instability, that results in large scale breaks and rearrangements (Kha *et al.*, 2010). Thus, in addition to serving as a sink for superhelical tension, there are several potential functions for Z-DNA that may be either beneficial or deleterious to the cell.

4.4 H-DNA: Three's a crowd

When a single DNA strand invades the major groove of a DNA duplex, a triple helical structure is generated (Fig. 9). In order for the duplex to accommodate this third strand, it must unwind to broaden the major groove; thus, such triple-stranded helices are favored in negatively supercoiled DNA (Mirkin, 2008). The invading third strand can be intermolecular or intramolecular.

The interaction between strands involve the Hoogsteen edge of the Watson-Crick base pairs (Fig. 3) of the duplex to form base triplets, leading to the name H-DNA for such triplex structures. H-DNA is formed primarily in mirror repeat sequences (sequences that have dyad symmetry within a strand, as in …AGAGGGnnnGGGAGA…, defined by the

sequence preference to form base triplets). Mirror-repeats occur randomly in prokaryotes, but are three to six times more frequent in eukaryotic genomes (Schroth and Ho, 1995). Specific H-DNA forming sequences have been identified in multiple promoter regions with documented effects on gene expression of several disease related genes, including c-myc (Kinniburgh, 1989) and c-Ki-ras (Pestov *et al.*, 1991). As with Z-DNA, the repeating sequence motif of H-DNA appears to be a source of genetic instability resulting from double-strand breaks. Wang and Vasquez (2004) reported a ~20 fold increase in mutation frequency upon incorporation of an H-DNA forming sequence found in the c-myc promoter region into mammalian cells. These results suggest that naturally occurring DNA sequences can cause increased mutagenesis via non-standard DNA structure formation.

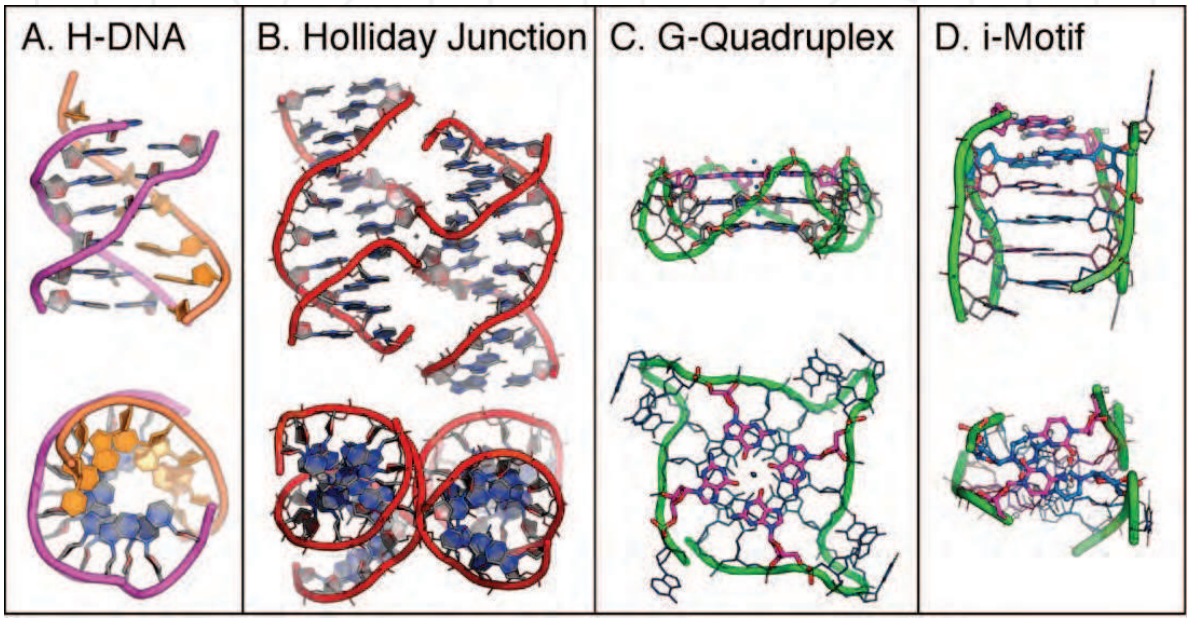

Fig. 9. Three- and four-stranded structures of DNA. The structures of triplex H-DNA (Radhakrishnan and Patel, 1993), the Holliday junction (Eichman et al, 2002), human telomeric G-quartet (Parkinson et al., 2002), and the i-motif (Weil et al., 1999), are viewed along (top) and down (bottom) their helical axes.

4.5 HJ, G, and I: The four-stranded DNAs

There are several conformations of DNA that can be assembled from four strands. The three structures discussed here show very different and unique helical forms, starting with a conformation that is most similar to standard B-DNA, and leading through forms that differ dramatically from the original Watson-Crick model (Fig. 9).

4.5.1 The four-stranded Holliday junction

Robin Holliday proposed in 1964 that a four-stranded junction would be involved as an intermediate to allow reciprocal exchange of genetic information through recombination across two homologous DNA duplexes (Holliday, 1964). These intermediates, now referred to as Holliday junctions, are essential to several cellular processes including recombination dependent DNA lesion repair, viral integration, restarting of stalled replication forks, and proper segregation of homologous chromosomes during meiosis *(Cox et al.,* 2000; Declais *et al.*, 2003; Dickman *et al.*, 2002; Haber and Heyer, 2001; Nunes-Duby *et al.*, 1987;

Subramaniam *et al.*, 2003). The structure of the Holliday junction has been the focus of intense biophysical studies for several decades (Lilley, 1999). Through a set of clever studies in which immobilized junctions are specifically cut by restriction enzymes or probed with fluorescent dyes, DNA junctions were shown to adopt either an extended open-X form under low-salt conditions or a more compact stacked-X conformation as the negatively charged phosphate backbone becomes shielded under high-salt conditions. In the stacked-X form, two continuous DNA strands are connected by two crossover strands, each forming a tight U-turn at the cross-over point, which restricts the migration of the junction. Single molecule studies have shown that migration requires a transition to the open-X structure (McKinney *et al.*, 2003), and that this is fairly rapid. As a result, enzymes that catalyze cellular processes that require junction migration (for example, during recombination dependent DNA repair by the RuvABC complex (Dickman *et al.*, 2002)) will recognize and bind the extended and topologically unrestrained open-X structure, while those that do not require junction migration (such as many resolving enzymes in recombination, including the resolvases from T4 and T7 (Biertumpfel *et al.*, 2007; Hadden *et al.*, 2007)) have active sites that bind to the topologically restrained stacked-X type structure.

Around the end of the 20th century, two groups almost simultaneously solved the single-crystal structures of the DNA Holliday junction (Ortiz-Lombardía *et al.*, 1999; Eichman *et al.*, 2000). Both structures strongly resembled the model derived from the solution studies (McKinney *et al.*, 2003), showing the junction to be essentially two B-DNA double-helices, with standard Watson-Crick type base pairs, linked by two crossing strands that connect the duplexes. A unique set of hydrogen bonds helps to stabilize the tight U-turns at the cross-over points (Eichman *et al.*, 2002), and impose a strong sequence dependence in the formation of Holliday junctions, with the inverted repeats GGTACC > GGCGCC > (GATATC = GGGCCC) in their stability as four-stranded stacked-X junctions (Hays *et al.*, 2005). In addition, the interactions define an ~40° angle relating the two linked duplexes—the structure of an asymmetric junction showed no interactions at the junction center, and an interduplex angle of ~60° (Khuu and Ho, 2009), similar to that determined in solution for analogous constructs (McKinney *et al.*, 2003). The structure of the junction has now been determined with the drug psoralen (Eichman *et al.*, 2001), methylated cytosines (Vargason and Ho, 2002), and various types of cations (Thorpe *et al.*, 2003), all showing effects on the detailed geometry of this four-stranded intermediate (Watson *et al.*, 2004). The effect of sequence on the formation and geometry of junctions lead to a model in which even non-sequence specific resolvases may show sequence preference, not as a result of any specific recognition motif between the protein and the DNA, but from the thermodynamic propensity of certain sequences to promote formation of the junction (Khuu, 2006).

In replication, Holliday junctions are essential intermediates in double-strand break repair (Cox *et al.*, 2000) in which RecA facilitates invasion of a single-strand into a homologous double-strand sequence, followed by junction migration and resolution by RuvABC (RecG). Homologous recombination also plays a crucial role in rescuing replication forks that stall because of DNA damage. Recombination proteins repair double-strand ends produced when a replication fork encounters a single-strand interruption and help reset replication at stalled forks by converting blocked replication forks into Holliday junctions. Thus, DNA junctions are involved in the repair of damaged DNAs both during and after replication.

4.5.2 G-Quadruplexes

The four-stranded structures assembled from guanine-rich sequences are called G-quadruplexes or G-quartets. Such sequences are found primarily in telomeric DNA repeats

(3'-overhangs at chromosome ends (Patel *et al.*, 2007)), but have recently been identified in various other central regions of the genome, including centrometric sequences (Brooks *et al.*, 2010) and in the immunoglobulin switch region. The strands are held together by pairing the Watson-Crick edge of each guanine with the Hoogsteen edge of an adjacent guanine, creating a cyclic arrangement of four guanines into G-tetrads. These tetrads are stacked with a right-handed helical twist, and are stabilized by monovalent cations (Na^+ or K^+) coordinated to the O2 oxygens of the guanines, and sandwiched between the base stacks.
G-quartets can be formed from the association of one, two, or four G-rich DNA strands with various topologies (Mirkin, 2008). Of these, the topologies that can be adopted by single-strands are perhaps most important for G-rich sequences at the 3'-ends (telomeric ends) of chromosomes (characterized as a single–stranded overhang of a guanine-rich sequence that assembles into a nucleo-protein structure). Such sequences have been shown to form G-quadruplex structures, from the DNA in the marconucleus of a ciliate (Mergny *et al.*, 2002) to the exceptionally stable G-quartet formed under physiological conditions by the human telomeric repeats ($(GGGTTA)_3GGG$) (Parkinson *et al.*, 2002). The telomer ends are replicated through the reverse transcriptase function of telomerase, which is itself a protein-RNA complex (Zakian, 2009). The precise length of each telomere controls the cell's ability to replicate, suggesting a regulatory role for their G-quadruplex structures. In normal cells, the length of the telomeric region is reduced during each round of replication until the Hayflick limit is reached, at which point the cell enters apoptosis (Zakian, 2009). The misregulation of telomerase activity can lead to immortality of cells and associated tumorogensis.

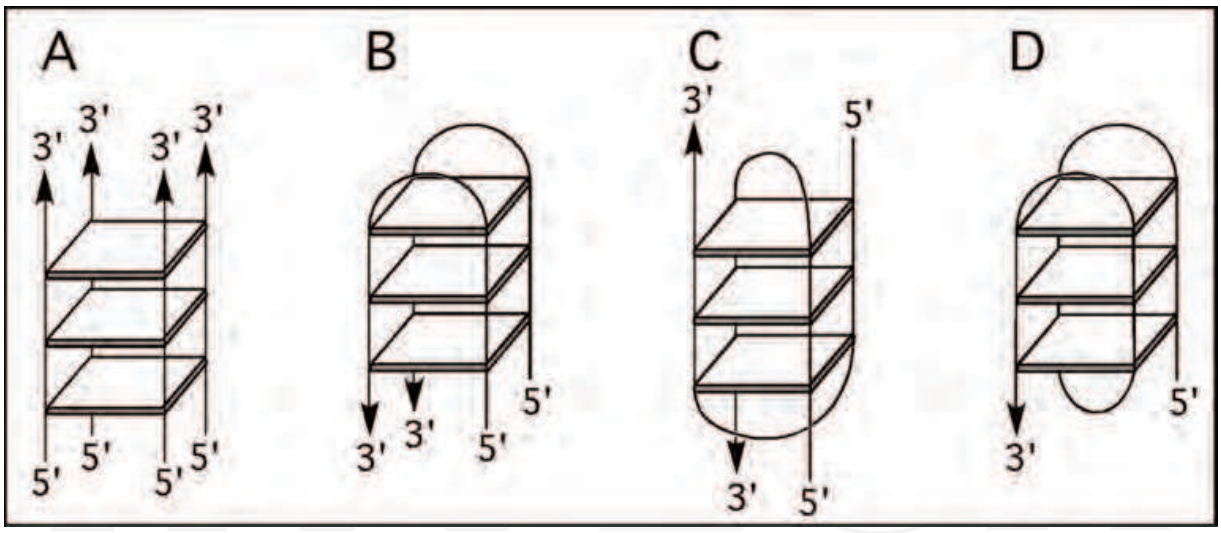

Fig. 10. Possible topologies for G-quartet structures. Topologies constructed from four parallel strands (A), from two strands that are non-crossing (B) or cross-over (C), and from a single strand (D).

Although it is easy to envision formation of a G-quartet structure at the single-stranded end of a chromosome, G-rich repeating sequences with the potential ability to form G-quadruplexes have also been identified at internal sites within genomes (Brooks *et al.*, 2010). Indeed, a recent study by Sarkies, *et al.* (Sarkies *et al.*, 2010) indicates that the specialized DNA polymerase Rev 1 is involved in replication through G-rich sequences and, when the polymerase is absent, DNA replication and histone recycling becomes uncoupled, leading to the assembly of nucleosomes with newly synthesized histones and, consequently, loss of epigenetic makers at or near these sites. Thus, internal G-quadruplex sequences are crucial for passing on to daughter cells genetic information beyond that of the linear sequence.

4.5.3 I-motifs

In order for a double-stranded G-rich region to extrude into a G-quartet structure, the complementary C-rich strand must also be extruded. The structure that is now associated with C-rich sequences is the four-stranded, intercalated i-motif. The i-motif, or I-form DNA, is fashioned from two parallel C-strands intercalated in a head-to-tail fashion [(Mills *et al.*, 2002). The two duplexes of poly(dC) are stabilized by base pairing the Watson-Crick edges of two cytosines to form hemi-protonated C•C+ pairs.

5. Getting from here to there: Structural transitions in DNA

B-DNA is recognized as the "standard" form in the cell; however, if everything remains standard and static, then life would not be as rich, nor might it exist at all. DNA is thus not only polymorphic, it is also dynamic. In this section, we will explore the mechanisms that drive DNA from the norm as B DNA, focusing on two transitions that present interesting and important insights into how DNA transforms between structural forms.

5.1 Going from B to A

As we have seen, A-type DNA plays an important role in replication as the induced form in the active site of DNA polymerase, allowing the non-sequence specific recognition of base mispairs in the template/daughter duplex. The transition from B- to A-DNA was one of the earliest characterized, with dehydration of DNA fibers showing a distinct shortening in the helical rise, unwinding of the helical twist, and broadening in the diameter (Franklin and Gosling, 1953a). The transition is also induced in solution by alcohol (a dehydrant), as well as methylation of cytosines (which affects the water structure around the base pairs). The question is, what are the structural and energetic steps involved in this transition? Although this is basically a transition from one right-handed antiparallel double-helix to another, several dramatic structural rearrangments must take place, including a conversion of the sugar pucker, along with large sliding and inclination of base pairs. The details of this conformational shift were observed crystallogaphically at the atomic level on the short DNA sequence GGCGCC (Vargason *et al.*, 2001), which was primarily in the B-form, but, upon cytosine methylation or bromination, adopts a number of conformational states, including true A-DNA forms and a set of logical intermediates between the B- and A-forms (Fig. 11). This study generates a structural map for how the sugar conformation works its way around the ring (Fig. 5), the order of translational and rotational distortions to the stacked base pairs, and the direction of propagation of a structural transition once initiated.

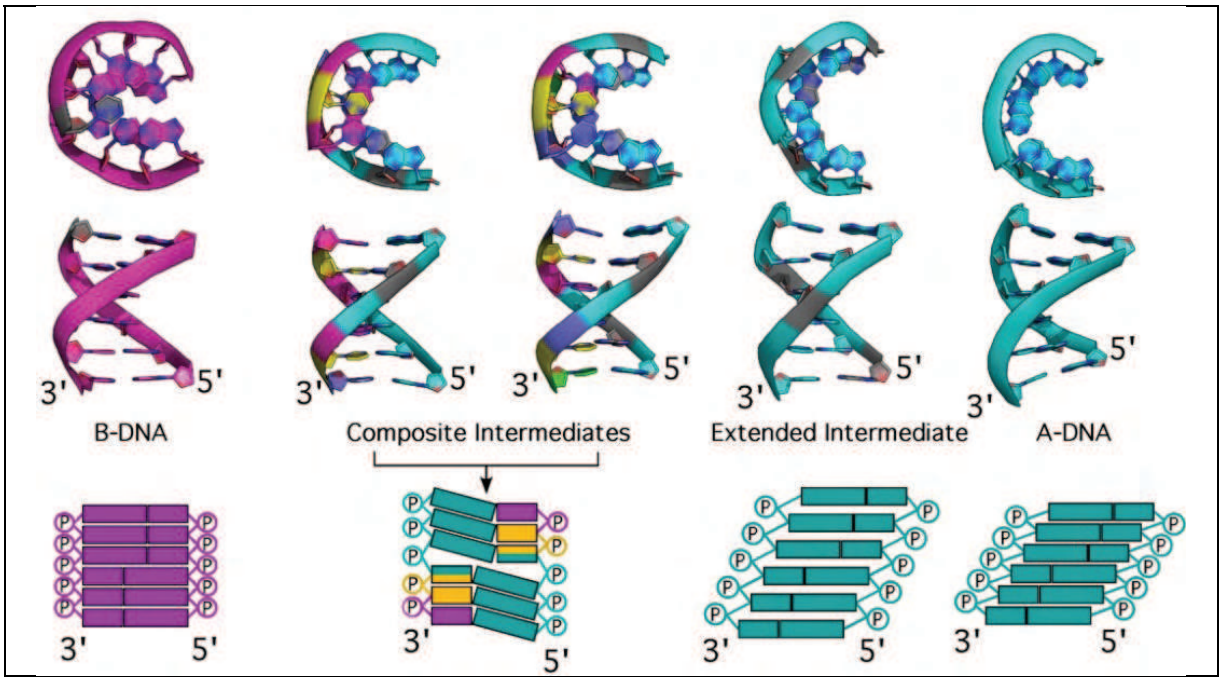

Fig. 11. B- to A-DNA transition. The structures of GGCGCC and methylated or brominated variants viewed down (top) and along (bottom) the helix axis. The series of structures show a transition from B-DNA, through a chimeric A-B intermediate and an extended intermediate, and leading finally to A-DNA. Nucleotides are colored according to their sugar puckers, as presented in Fig. 5.

The transition involves conversion of the sugar from the B-DNA C2′-*endo* pucker to C1′-*exo*, then O4′-*endo*, followed by C4′-*exo*, and finally to the C3′-*endo* pucker of A-DNA (Fig. 5) (Vargason *et al.*, 2001). Applying *ab initio* calculations on models of the deoxyribose derived from this study, we found that there is an ~4 kcal/mol energy barrier (primarily bonding energy) at the O4′-*endo* intermediate step. This is lower than the ~5-6 kcal/mol estimated for planar intermediates required for a direct conversion from C2′- to C3′-*endo*, and is similar to estimates from experimental (Olson and Sussman, 1982) and other *ab initio* calculations (Foloppe *et al.*, 2001) on the barrier (although about 2-fold higher than molecular dynamics estimates (Arora and Schlick, 2003; Harvey and Prabhakaran, 1986)).
Associated with the changes in sugar pucker are perturbations to the base stacking. As the sugars go through a transition from B- towards A-type sugars, the B-A chimeric intermediate (which is half B- and half A-type along each strand) induces a large buckle in the base pairs at the point of transition, which partially unstacks one of the two bases of the pair. The unstacking becomes complete when the sugars assume the full A-type pucker, resulting in an ~10% extension of the spacing between bases, or a rise of ~3.7 Å (Vargason *et al.*, 2000), thereby allowing the large slide and subsequent displacement of the base pairs away from the helical axis that is characteristic of A-DNA. Thus, large shifts between base pairs are predicated on breaking the base stacking interactions, as one would expect. In addition, it shows the transition to A-DNA propagating back towards the 5′-end of each strand. The tilt and roll that causes the inclination and resulting shortened rise of A-DNA are the final steps. The B- to A-DNA transition is unique in that specific intermediates have been trapped to provide an atomic level map for the transition—this is perhaps the most detailed description of a complete structural transition of any biological macromolecule.

5.2 Switching hands: The B- to Z-DNA transition

A more dramatic transition is from the right-handed B- to left-handed Z-DNA (Fig. 12), which has been studied extensively in solution and in plasmids. The B-Z transition, however, does not simply twist a right-handed double-helix in the opposite direction. The sugar for alternating nucleotides along a strand change from C2′-*endo* to C3′-*endo* puckers, concommitant with rotation of the base from the *anti*- to the *syn*-conformations. More significantly, the "sense" of the duplex must change—*i.e.,* the direction of the major and minor grooves are swapped (Dickerson, 1992).

In order to accommodate all of these radical changes, there is a junction with an overall zero twist (the B-Z junction) that serves to splice the right- and left-handed twisted duplexes (Peck and Wang, 1983). The structure of this junction was determined in a clever way using a Z-DNA binding protein to stabilize half the DNA in the left-handed form, while allowing the other half to remain in its relaxed B-form (Ha *et al.*, 2005). The structure shows that the bases at the B-Z junction itself have flipped out, which would allow for transition of the sugar pucker and rotation of the bases. It also allows the bases, when they pair again, to change the direction of the grooves sense, while maintaining stacking between the left- and right-handed columns. The B-Z transition, therefore, can be thought of as initiating with a melting of two base pairs (two B-Z junctions, with a nucleation energy of ~10 kcal/mol (Peck and Wang, 1983)), with each junction subsequently migrating in opposite directions to allow the propagation of the left-handed DNA between them (the propagation energy per base pair being sequence dependent and lowest in alternating GC dinucleotides (Ellison *et al.*, 1985)).

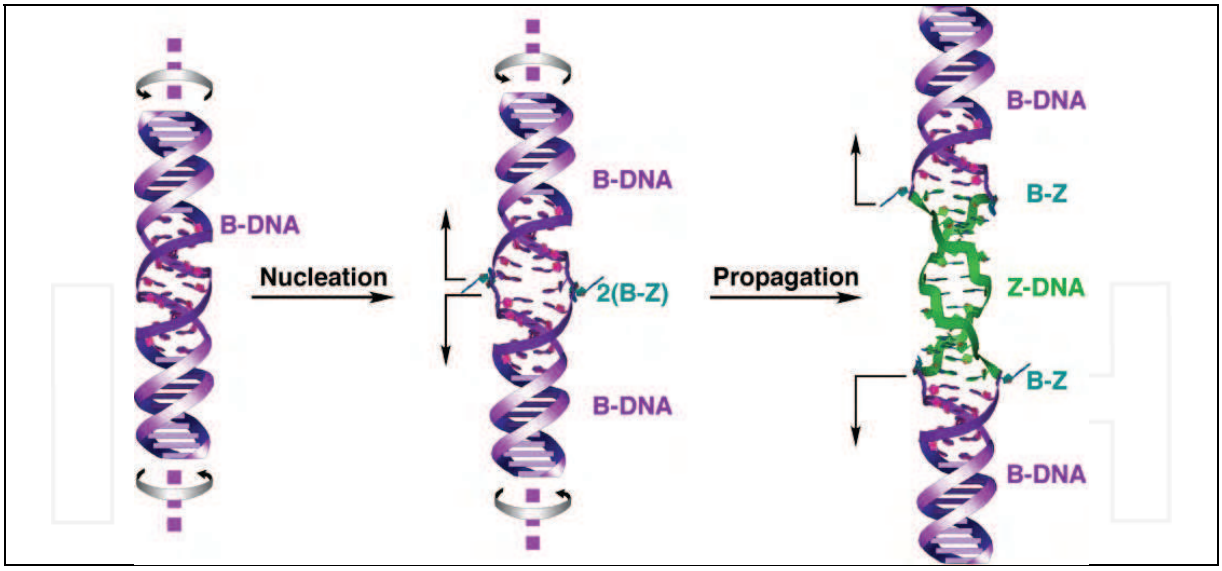

Fig. 12. B- to Z-DNA transition. B-DNA, when unwound by negative supercoiling, will first extrude two flipped out base pairs (serving as two B-Z junctions). Further unwinding results in the formation of left-handed Z-DNA between the two junctions as they migrate in opposite directions.

6. Conclusion

In this review, we have discussed a plethora of structures that come from physical biochemical studies, and show how these structures are defined by sequence and how they transform. Through its history, there has always been a nagging question of "Is this structure relevant?" Clearly, the B-DNA double-helix is relevant, not only to replication, but also to nearly all genetic processes. However, a clearer understanding for the biological roles of the non-B-type DNAs will require a detailed mapping of such structures (Ho, 2009), either experimentally or computationally, across genomes from various organisms.

7. References

Allemand, J.F., Bensimon, D., Lavery, R., & Croquette, V. (1998). Stretched and overwound DNA forms a Pauling-like structure with exposed bases. *Proc Natl Acad Sci, USA* 95: 14152-14157.

Arora, K., & Schlick, T. (2003). Deoxyadenosine sugar puckering pathway simulated by the stochastic difference equation algorithm. *Chem Phys Lett* 378: 1-8.

Ausio, J., Zhou, G., & van Holde, K. (1987). A reexamination of the reported B----Z DNA transition in nucleosomes reconstituted with poly(dG-m5dC).poly(dG-m5dC). *Biochemistry* 26: 5595-5599.

Avery, O.T., Macleod, C.M., & McCarty, M. (1944). Studies on the chemical nature of the substance inducing transformation of pneumococcal types: Induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus Type Iii. *J Exp Med* 79: 137-158.

Biertumpfel, C., Yang, W., & Suck, D. (2007). Crystal structure of T4 endonuclease VII resolving a Holliday junction. *Nature* 449: 616-620.

Brooks, T.A., Kendrick, S., & Hurley, L. Making sense of G-quadruplex and i-motif functions in oncogene promoters. *FEBS J* 277: 3459-3469.

Burley, S.K. (1996). X-ray crystallographic studies of eukaryotic transcription initiation factors. *Philos Trans R Soc Lond B Biol Sci* 351: 483-489.

Champ, P.C., Maurice, S., Vargason, J.M., Camp, T., & Ho, P.S. (2004). Distributions of Z-DNA and nuclear factor I in human chromosome 22: a model for coupled transcriptional regulation. *Nucleic Acids Res* 32: 6501-6510.

Chargaff, E. (1950). Chemical specificity of nucleic acids and mechanism of their enzymatic degradation. *Experientia* 6: 201-209.

Coll, M., Frederick, C.A., Wang, A.H., & Rich, A. (1987). A bifurcated hydrogen-bonded conformation in the d(A.T) base pairs of the DNA dodecamer d(CGCAAATTTGCG) and its complex with distamycin. *Proc Natl Acad Sci, USA* 84: 8385-8389.

Cox, M.M., Goodman, M.F., Kreuzer, K.N., Sherratt, D.J., Sandler, S.J., & Marians, K.J. (2000). The importance of repairing stalled replication forks. *Nature* 404: 37-41.

Day, R.O., Seeman, N.C., Rosenberg, J.M., & Rich, A. (1973). A crystalline fragment of the double helix: the structure of the dinucleoside phosphate guanylyl-3',5'-cytidine. *Proc Natl Acad Sci U S A* 70: 849-853.

Declais, A.C., Fogg, J.M., Freeman, A.D., Coste, F., Hadden, J.M., Phillips, S.E., & Lilley, D.M. (2003). The complex between a four-way DNA junction and T7 endonuclease I. *EMBO J* 22: 1398-1409.

Dickerson, R.E. (1992). DNA structure from A to Z. *Methods Enzymol* 211: 67-111.

Dickerson, R.E. (1999). Helix structure and molecular recognition by B-DNA. In *Oxford Handbook of Nucleic Acid Structure*, S. Neidle, ed. (New York, Oxford University Press), pp. 145-197.

Dickerson, R.E., Goodsell, D.S., & Neidle, S. (1994). "...the tyranny of the lattice...". *Proc Natl Acad Sci, USA* 91: 3579-3583.

Dickman, M.J., Ingleston, S.M., Sedelnikova, S.E., Rafferty, J.B., Lloyd, R.G., Grasby, J.A., & Hornby, D.P. (2002). The RuvABC resolvasome. *Eur J Biochem* 269: 5492-5501.

Drew, H.R., Wing, R.M., Takano, T., Broka, C., Tanaka, S., Itakura, K., & Dickerson, R.E. (1981). Structure of a B-DNA dodecamer: conformation and dynamics. *Proc Natl Acad Sci USA* 78: 2179-2183.

Eichman, B.F., Mooers, B.H.M., Alberti, M., Hearst, J.E., & Ho, P.S. (2001). The crystal structures of psoralen cross-linked DNAs: drug dependent formation of Holliday junctions. *J Mol Biol* 301: 15-26.

Eichman, B.F., Ortiz-Lombardia, M., Aymami, J., Coll, M., & Ho, P.S. (2002). The inherent properties of DNA four-way junctions: comparing the crystal structures of holliday junctions. *J Mol Biol* 320: 1037-1051.

Eichman, B.F., Vargason, J.M., Mooers, B.H.M., & Ho, P.S. (2000). The Holliday junction in an inverted repeat sequence: sequence effects on the structure of four-way junctions. *Proc Natl Acad Sci, USA* 97: 3971-3976.

Ellison, M.J., Kelleher, R.J., 3rd, Wang, A.H., Habener, J.F., & Rich, A. (1985). Sequence-dependent energetics of the B-Z transition in supercoiled DNA containing nonalternating purine-pyrimidine sequences. *Proc Natl Acad Sci, USA* 82: 8320-8324.

Foloppe, N., Nilsson, L., & MacKerell, A.D., Jr. (2001). Ab initio conformational analysis of nucleic acid components: intrinsic energetic contributions to nucleic acid structure and dynamics. *Biopolymers* 61: 61-76.

Franklin, R.E., & Gosling, R.G. (1953a). Evidence for 2-chain helix in crystalline structure of sodium deoxyribonucleate. *Nature* 172: 156-157.

Franklin, R.E., & Gosling, R.G. (1953b). Molecular Configuration in Sodium Thymonucleate. *Nature (London)* 171: 740-741.

Ghosh, A., & Bansal, M. (2003). A glossary of DNA structures from A to Z. *Acta Crystallogr D Biol Crystallogr* 59: 620-626.

Greenbaum, J.A., Pang, B., & Tullius, T.D. (2007). Construction of a genome-scale structural map at single-nucleotide resolution. *Genome Res* 17: 947-953.

Ha, S.C., Lowenhaupt, K., Rich, A., Kim, Y.G., & Kim, K.K. (2005). Crystal structure of a junction between B-DNA and Z-DNA reveals two extruded bases. *Nature* 437: 1183-1186.

Haber, J.E., & Heyer, W.D. (2001). The fuss about Mus81. *Cell* 107: 551-554.

Hadden, J.M., Declais, A.C., Carr, S.B., Lilley, D.M., & Phillips, S.E. (2007). The structural basis of Holliday junction resolution by T7 endonuclease I. *Nature* 449: 621-624.

Hamelberg, D., Williams, L.D., & Wilson, W.D. (2001). Influence of the dynamic positions of cations on the structure of the DNA minor groove: sequence-dependent effects. *J Am Chem Soc* 123: 7745-7755.

Harvey, S.C., & Prabhakaran, M. (1986). Ribose Puckering: Structure, Dynamics, Energetics, and the Pseudorotation Cycle. *J Amer Chem Soc* 108: 6128-6136

Hays, F.A., Teegarden, A., Jones, Z.J., Harms, M., Raup, D., Watson, J., Cavaliere, E., & Ho, P.S. (2005). How sequence defines structure: a crystallographic map of DNA structure and conformation. *Proc Natl Acad Sci, USA* 102: 7157-7162.

Hendrickson, W.A., Smith, J.L., Phizackerley, R.P., & Merritt, E.A. (1988). Crystallographic structure analysis of lamprey hemoglobin from anomalous dispersion of synchrotron radiation. *Proteins* 4: 77-88.

Hershey, A.D., & Chase, M. (1952). Independent functions of viral protein and nucleic acid in growth of bacteriophage. *J Gen Physiol* 36: 39-56.

Ho, P.S. (2009). Thermogenomics: thermodynamic-based approaches to genomic analyses of DNA structure. *Methods* 47: 159-167.

Ho, P.S., Frederick, C.A., Quigley, G.J., van der Marel, G.A., van Boom, J.H., Wang, A.H., & Rich, A. (1985). G.T wobble base-pairing in Z-DNA at 1.0 A atomic resolution: the crystal structure of d(CGCGTG). *EMBO J* 4: 3617-3623.

Holliday, R. (1964). A mechanism for gene conversion in fungi. *Genet Res* 5: 282-304.

Kha, D.T., Wang, G., Natrajan, N., Harrison, L., & Vasquez, K.M. Pathways for double-strand break repair in genetically unstable Z-DNA-forming sequences. *J Mol Biol* 398: 471-480.

Khuu, P., & Ho, P.S. (2009). A rare nucleotide tautomer in the structure of an asymmetric DNA junction. *Biochemistry* 48: 7824-7832.

Khuu, P., Sandor, M., DeYoung, J., & Ho, P.S. (2007). Phylogenomic analysis of the emergence of GC-rich transcription elements. *Proc Natl Acad Sci, USA* 104: 16528-16533.

Khuu, P.A., Voth, A. R., Hays, F. A. and Ho, P. S. (2006). The stacked-X DNA Holliday junction and protein recognition. *J Mol Recognit* 19: 1-9.

Kiefer, J.R., Mao, C., Braman, J.C., & Beese, L.S. (1998). Visualizing DNA replication in a catalytically active Bacillus DNA polymerase crystal. *Nature* 391: 304-307.

Kinniburgh, A.J. (1989). A cis-acting transcription element of the c-myc gene can assume an H-DNA conformation. *Nucleic Acids Res* 17: 7771-7778.

Klotz, I.M. & Franzen, J.S. (1962). Hydrogen bonds between model peptide groups in solution. *J Am Chem Soc* 84: 3461-3466.

Kool, E.T. (2001). Hydrogen bonding, base stacking, and steric effects in dna replication. *Annu Rev Biophys Biomol Struct* 30: 1-22.

Lavery, R., & Zakrzewska, K (1998) in *Oxford Handbook of Nucleic Acid Structure*, S. Neidle, ed. (New York, Oxford University Press), pp39-76.

Liepinsh, E., Otting, G., & Wüthrich, K. (1992). NMR observation of individual molecules of hydration water bound to DNA duplexes: direct evidence for a spine of hydration water present in aqueous solution. *Nucleic Acids Res* 20: 6549-6553.

Lilley, D.M.J. (1999). Structures and interactions of helical junctions in nucleic acids. In *Oxford Handbook of Nucleic Acid Structure*, S. Neidle, ed. (New York, Oxford University Press), pp. 471-498.

Liu, L.F., & Wang, J.C. (1987). Supercoiling of the DNA template during transcription. *Proc Proc Natl Acad Sci, USA* 84: 7024-7027.

Liu, R., Liu, H., Chen, X., Kirby, M., Brown, P.O., & Zhao, K. (2001). Regulation of CSF1 promoter by the SWI/SNF-like BAF complex. *Cell* 106: 309-318.

Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., & Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* 389: 251-260.

Manning, G.S. (1977). Limiting laws and counterion condensation in polyelectrolyte solutions. IV. The approach to the limit and the extraordinary stability of the charge fraction. *Biophys Chem* 7: 95-102.

McKinney, S.A., Declais, A.C., Lilley, D.M., & Ha, T. (2003). Structural dynamics of individual Holliday junctions. *Nat Struct Biol* 10: 93-97.

Mergny, J.L., Riou, J.F., Mailliet, P., Teulade-Fichou, M.P., & Gilson, E. (2002). Natural and pharmacological regulation of telomerase. *Nucleic Acids Res* 30: 839-865.

Meselson, M., & Stahl, F.W. (1958). The replication of DNA in Escherichia coli. *Proc Natl Acad Sci, USA* 44: 671-682.

Mills, M., Lacroix, L., Arimondo, P.B., Leroy, J.L., Francois, J.C., Klump, H., & Mergny, J.L. (2002). Unusual DNA conformations: implications for telomeres. *Curr Med Chem Anticancer Agents* 2: 627-644.

Mirkin, S.M. (2008). Discovery of alternative DNA structures: a heroic decade (1979-1989). *Front Biosci* 13: 1064-1071.

Neidle, S. (1999). Oxford Handbook of Nucleic Acid Structure (Oxford, Oxford Press).

Ng, H.L., Kopka, M.L., & Dickerson, R.E. (2000). The structure of a stable intermediate in the A <--> B DNA helix transition. *Proc Natl Acad Sci, USA* 97: 2035-2039.

Nunes-Duby, S.E., Matsumoto, L., & Landy, A. (1987). Site-specific recombination intermediates trapped with suicide substrates. *Cell* 50: 779-788.

Olson, W.K., & Sussman, J.L. (1982). How flexible is the furanose ring? 1. A comparison of experimental and theoretical studies. *J Am Chem Soc*: 270-278.

Ortiz-Lombardía, M., González, A., Eritja, R., Aymamí, J., Azorín, F., & Coll, M. (1999). Crystal structure of a DNA Holliday junction. *Nat Struct Biol* 6: 913-917.

Parkinson, G.N., Lee, M.P., & Neidle, S. (2002). Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature* 417: 876-880.

Patel, D.J., Phan, A.T., & Kuryavyi, V. (2007). Human telomere, oncogenic promoter and 5'-UTR G-quadruplexes: diverse higher order DNA and RNA targets for cancer therapeutics. *Nucleic Acids Res* 35: 7429-7455.

Pauling, L., & Corey, R.B. (1953). A proposed structure for the nucleic acids. *Proc Natl Acad Sci, USA* 39: 84-97.

Peck, L.J., & Wang, J.C. (1983). Energetics of B-to-Z transition in DNA. *Proc Natl Acad Sci, USA* 80: 6206-6210.

Pestov, D.G., Dayn, A., Siyanova, E., George, D.L., & Mirkin, S.M. (1991). H-DNA and Z-DNA in the mouse c-Ki-ras promoter. *Nucleic Acids Res* 19: 6527-6532.

Pohl, F.M., & Jovin, T.M. (1972). Salt-induced co-operative conformational change of a synthetic DNA: equilibrium and kinetic studies with poly (dG-dC). *J Mol Biol* 67: 375-396.

Privé, G.G., Yanagi, K., & Dickerson, R.E. (1991). Structure of the B-DNA decamer C-C-A-A-C-G-T-T-G-G and comparison with isomorphous decamers C-C-A-A-G-A-T-T-G-G and C-C-A-G-G-C-C-T-G-G. *J Mol Biol* 217: 177-199.

Radhakrishnan, I., & Patel, D.J. (1993). Solution structure of a purine.purine.pyrimidine DNA triplex containing G.GC and T.AT triples. *Structure* 1: 135-152.

Rich, A., & Zhang, S. (2003). Timeline: Z-DNA: the long road to biological function. *Nat Rev Genet* 4: 566-572.

Rohs, R., West, S.M., Sosinsky, A., Liu, P., Mann, R.S., & Honig, B. (2009). The role of DNA shape in protein-DNA recognition. *Nature* 461: 1248-1253.

Rosenberg, J.M., Seeman, N.C., Kim, J.J., Suddath, F.L., Nicholas, H.B., & Rich, A. (1973). Double helix at atomic resolution. *Nature* 243: 150-154.

Saenger, W. (1984) *Principles of Nucleic Acid Structure*, Spinger-Verlag, New York

Sarkies, P., Reams, C., Simpson, L.J., & Sale, J.E. Epigenetic instability due to defective replication of structured DNA. *Mol Cell* 40: 703-713.

Schroth, G.P., Chou, P.J., & Ho, P.S. (1992). Mapping Z-DNA in the human genome. Computer-aided mapping reveals a nonrandom distribution of potential Z-DNA-forming sequences in human genes. *J Biol Chem* 267: 11846-11855.

Schroth, G.P., & Ho, P.S. (1995). Occurrence of potential cruciform and H-DNA forming sequences in genomic DNA. *Nucleic Acids Res* 23: 1977-1983.

Subramaniam, S., Tewari, A.K., Nunes-Duby, S.E., & Foster, M.P. (2003). Dynamics and DNA substrate recognition by the catalytic domain of lambda integrase. *J Mol Biol* 329: 423-439.

Thorpe, J.H., Gale, B.C., Teixeira, S.C., & Cardin, C.J. (2003). Conformational and hydration effects of site-selective sodium, calcium and strontium ion binding to the DNA Holliday junction structure d(TCGGTACCGA)(4). *J Mol Biol* 327: 97-109.

Vargason, J.M., Eichman, B.F., & Ho, P.S. (2000). The extended and eccentric E-DNA structure induced by cytosine methylation or bromination. *Nat Struct Biol* 7: 758-761.

Vargason, J.M., Henderson, K., & Ho, P.S. (2001). A crystallographic map of the transition from B-DNA to A-DNA. *Proc Natl Acad Sci, USA* 98: 7265-7270.

Vargason, J.M., & Ho, P.S. (2002). The effect of cytosine methylation on the structure and geometry of the Holliday junction: the structure of d(CCGGTACm5CGG) at 1.5 A resolution. *J Biol Chem* 277: 21041-21049.

Wang, A.H., Quigley, G.J., Kolpak, F.J., Crawford, J.L., van Boom, J.H., van der Marel, G., & Rich, A. (1979). Molecular structure of a left-handed double helical DNA fragment at atomic resolution. *Nature* 282: 680-686.

Wang, G., & Vasquez, K.M. (2004). Naturally occurring H-DNA-forming sequences are mutagenic in mammalian cells. *Proc Natl Acad Sci, USA* 101: 13448-13453.

Watson, J., Hays, F.A., & Ho, P.S. (2004). Definitions and analysis of DNA Holliday junction geometry. *Nucleic Acids Res* 32: 3017-3027.

Watson, J.D. (1968). The Double Helix: A Personal Account of the Discovery of the Structure of DNA (Atheneum).

Watson, J.D., & Crick, F.H. (1953a). Genetical implications of the structure of deoxyribonucleic acid. *Nature* 171: 964-967.

Watson, J.D., & Crick, F.H. (1953b). Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* 171: 737-738.

Weil, J., Min, T., Yang, C., Wang, S., Sutherland, C., Sinha, N., & Kang, C. (1999). Stabilization of the i-motif by intramolecular adenine-adenine-thymine base triple in the structure of d(ACCCT). *Acta Crystallogr D Biol Crystallogr* 55: 422-429.

Zakian, V.A. (2009). The ends have arrived. *Cell* 139: 1038-1040.

DNA Replication-Current Advances

Edited by Dr Herve Seligmann

ISBN 978-953-307-593-8
Hard cover, 694 pages
Publisher InTech
Published online 01, August, 2011
Published in print edition August, 2011

The study of DNA advanced human knowledge in a way comparable to the major theories in physics, surpassed only by discoveries such as fire or the number zero. However, it also created conceptual shortcuts, beliefs and misunderstandings that obscure the natural phenomena, hindering its better understanding. The deep conviction that no human knowledge is perfect, but only perfectible, should function as a fair safeguard against scientific dogmatism and enable open discussion. With this aim, this book will offer to its readers 30 chapters on current trends in the field of DNA replication. As several contributions in this book show, the study of DNA will continue for a while to be a leading front of scientific activities.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

P. Shing Ho and Megan Carter (2011). DNA Structure: Alphabet Soup for the Cellular Soul, DNA Replication-Current Advances, Dr Herve Seligmann (Ed.), ISBN: 978-953-307-593-8, InTech, Available from: http://www.intechopen.com/books/dna-replication-current-advances/dna-structure-alphabet-soup-for-the-cellular-soul

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821