CAMBRIDGE

Cambridge University Press
978-1-107-03493-8 - Biophysics of DNA
Alexander Vologodskii
Excerpt
More information

1 DNA structures

1.1 Chemical structure and conformational flexibility of single-stranded DNA

Single-stranded DNA (ssDNA) is the building base for the double helix and other DNA structures. All these structures are formed due to noncovalent interactions between the components of ssDNA. ssDNA consists of a backbone of repeating units and bases that are attached to each unit as side chains (Fig. 1.1).

An isolated part of the repeating unit that consists of a base and the sugar is called a *nucleoside*. If a phosphate group is added to the nucleoside, it becomes a 3′- or 5′-*nucleotide*, depending on where the phosphate group is bound. Each repeating unit of the backbone consists of sugar and phosphate and has six skeletal bonds. The backbone has clear directionality, and the method of numbering of carbon atoms of the sugar, shown in Fig. 1.1, identifies 3′–5′ or 5′–3′ directions. It is common to assume a 5′–3′ direction of the polynucleotide chain when presenting a sequence of bases.

There is an important degree of freedom in isolated nucleosides that is related to rotation around the bond connecting the sugar and a base, the β-glycosidic bond. The rotation angle, χ, is measured with reference to the orientation of $O^{1'}$–$C^{1'}$ and N^9–C^8 bonds for purines and to the orientation of $O^{1'}$–$C^{1'}$ and N^1–C^6 bonds for pyrimidines. Although many different values of χ are sterically allowed, two major rotational isomers, called *anti* and *syn*, are particularly important. For the *anti* conformation χ is close to 0°, and for *syn* χ is around 210°. The conformations are diagramed in Fig. 1.2.

The bond lengths and the angles between the adjacent bonds do not change notably. The remarkable conformational flexibility of ssDNA is due to six rotation angles in each repeating unit of the backbone. Of course, the rotation angles cannot accept just any values, since there are many potential steric clashes between chemical groups of the unit.

The deoxyribose ring is also able to adopt different conformations. Four of the five atoms of the ring tend to be in the plane, while the fifth is out of the plane. Depending on which atom is out of the plane, $C^{2'}$ or $C^{3'}$, the ring can be in one of four conformations. In the $C^{2'}$-*endo* conformation $C^{2'}$ lies above the plane, on the same side as the base and 5′-carbon (Fig. 1.3). The $C^{3'}$-*exo* conformation is similar to $C^{2'}$-*endo*, although the $C^{3'}$ atom deviates from the plane more than $C^{2'}$. Correspondingly, in the $C^{3'}$-*endo* conformation $C^{3'}$ is located above the ring, and the $C^{2'}$-*exo* conformation is similar to $C^{3'}$-*endo*.

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-1-107-03493-8 - Biophysics of DNA
Alexander Vologodskii
Excerpt
More information

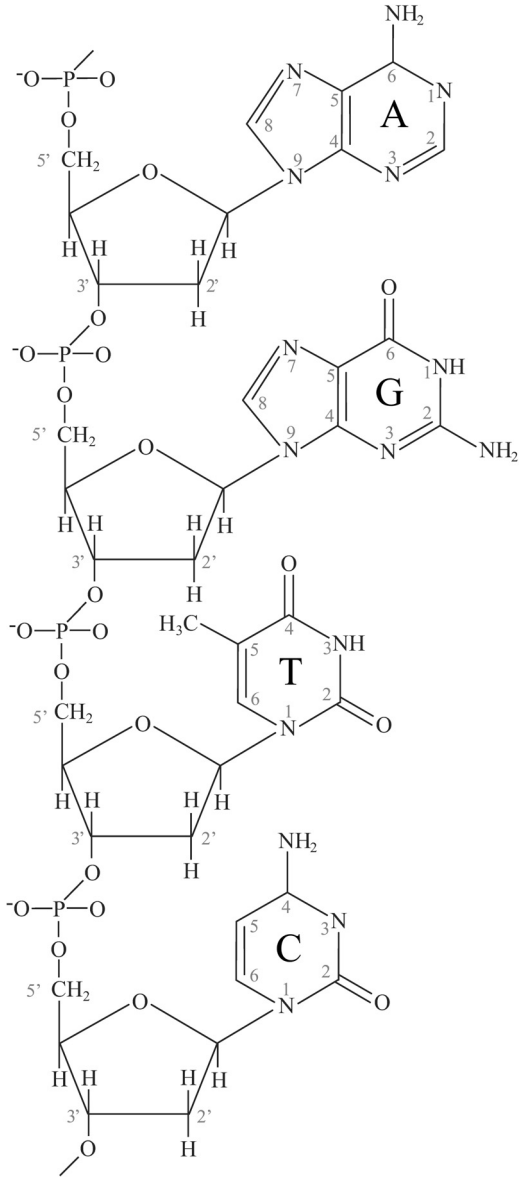

Figure 1.1 Chemical structure of ssDNA. Each repeating unit of the molecule backbone consists of 2′-deoxyribose and a negatively charged phosphate residue. The side chains are represented by one of four bases: adenine (A), guanine (G), thymine (T) or cytosine (C). Adenine and guanine belong to the group of purines, while the smaller thymine and cytosine are pyrimidines. The numbering systems of carbon and nitrogen atoms in the backbone and the bases are shown by gray digits.

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-1-107-03493-8 - Biophysics of DNA
Alexander Vologodskii
Excerpt
More information

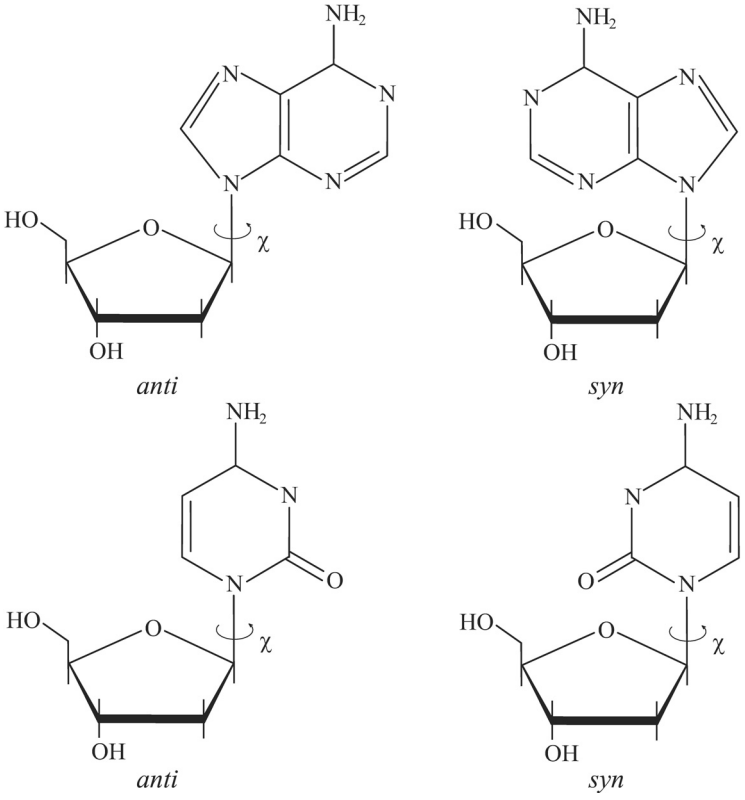

Figure 1.2 *Anti* and *syn* conformations of purines and pyrimidines. The actual conformations differ by 210° in angle χ and are only approximately represented in the figure.

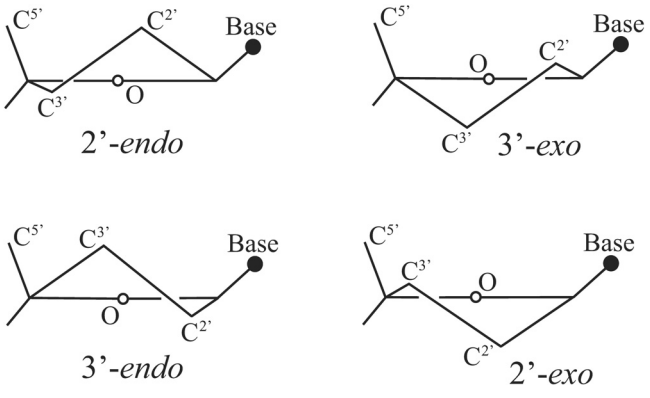

Figure 1.3 Major conformations of the deoxyribose ring. The plane containing O, $C^{1'}$ and $C^{4'}$ is perpendicular to the page.

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-1-107-03493-8 - Biophysics of DNA
Alexander Vologodskii
Excerpt
More information

The nucleosides are uncharged at pH values between 3 and 9, while the phosphate residues are negatively charged. Thus, each nucleotide of ssDNA has a single negative charge, with the exception of the end of the polynucleotide chain, where the phosphate group has a double negative charge at pH 7.

1.2 Double helices

The most important structural form of DNA is called the B form. It is this form of DNA, discovered by Watson and Crick in 1953 (Watson & Crick 1953), that started modern biology. It became clear later that DNA can adopt other double-stranded forms as well. All double-stranded forms share the most important structural feature, the base pairing. The base pairs are formed by bases from opposite strands. There are two canonical base pairs, AT and GC. Thus, a base in one strand determines the base at the corresponding position in the other strand. This means that the sequence of bases in one strand completely determines the sequence in the other strand. This is a key property for DNA biological functioning.

There are two hydrogen bonds in an AT base pair and three in a GC pair (Fig. 1.4). This difference in the number of hydrogen bonds is responsible for a higher thermal stability of GC base pairs (although base pairing per se does not stabilize the double helix (Yakovchuk *et al.* 2006)). The base pairs do not just have very close dimensions, but also their external geometries related to the backbones are nearly identical and have an important pseudosymmetry. The bonds between the nitrogens of the bases and $C^{1'}$ atoms of the sugar rings form the same angle of 51.5° with the $C^{1'}$–$C^{1'}$ line for all four bases of AT and GC base pairs (see Fig. 1.4). Base-pair geometry specifies the distance between $C^{1'}$ atoms across the pair, and this distance is exactly the same for both AT and GC pairs. Therefore, in either of two orientations (AT or TA, and GC or CG) the base pairs can be very well incorporated in a uniform helical structure of the backbones. It is this pseudosymmetry that makes Watson–Crick base pairing so unique. Many other possible patterns of base pairing do not have this symmetry and cannot be incorporated into a uniform structure of the backbones.

1.2.1 B-DNA

B-DNA, a form that the double helix has in aqueous solutions, is a right-handed helix with a helical period close to 10.5 bp per turn (Wang 1979, Peck & Wang 1981, Goulet *et al.* 1987). Its external diameter is approximately 2.0 nm (Dickerson & Ng 2001). The complementary strands have antiparallel orientations. The base pairs are located inside the helix, while the backbones are at the helix exterior (Fig. 1.5). The helix has two dyad symmetry axes that are perpendicular to the helix axis (assuming that the helix ends are extended to infinity). One of them passes through the plane of the bases (see Fig. 1.4). The other one passes between two base pairs. These are true symmetry axes for the backbone and pseudosymmetry axes for the base pairs. For

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-1-107-03493-8 - Biophysics of DNA
Alexander Vologodskii
Excerpt
More information

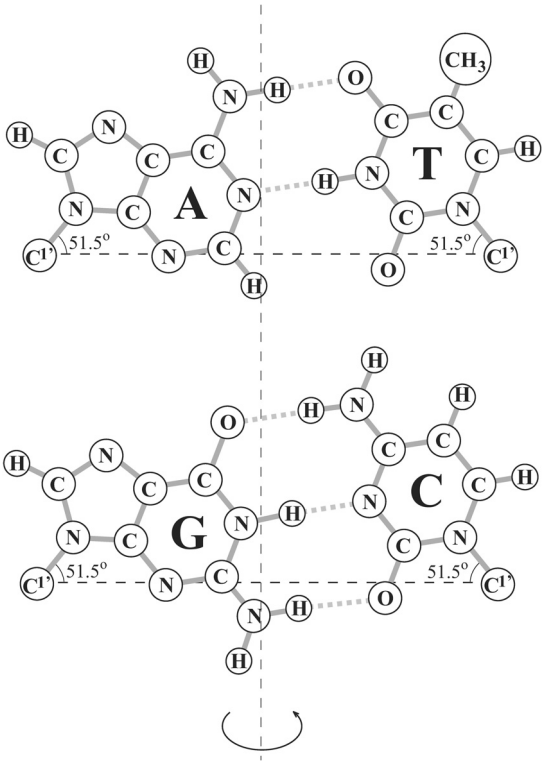

Figure 1.4 Complementary base pairs of DNA. The bonds directed to $C^{1'}$ atoms of the deoxyribose from the nitrogens of the bases form the same angles with the $C^{1'}$–$C^{1'}$ line. The dyad symmetry axis is shown by the dashed line. The distance between $C^{1'}$ atoms is equal to 1.085 nm for both base pairs.

the double helix formed by self-complementary polynucleotides poly(AT)poly(AT) and poly(GC)poly(GC), however, the latter axis is a true symmetry axis for the entire helix.

The average planes of adjacent base pairs are nearly perpendicular to the helix axis, although the base pairs are not flat. In B-DNA they form a propeller with 16° between the planes of the bases. The planes are separated by 0.34 nm. There is a strong interaction between the bases of adjacent base pairs, called the stacking interaction. This is the most important interaction stabilizing the double helix.

The helix has two grooves, the minor and major ones. The bases are more exposed in the major groove (see Fig. 5.5). The major groove is approximately 2.2 nm in width, while the width of the minor groove is close to 1.2 nm (Wing *et al.* 1980). The glycosyl angles in B-DNA correspond to *anti* conformations. The sugar ring has $C^{2'}$-*endo* conformation.

Structural parameters of B-DNA weakly depend on the sequence, although precise information on this dependence is limited. There are an enormous number of X-ray data on structures of oligonucleotide duplexes (Berman *et al.* 1992), but these structural data are affected by the forces of crystal packing. Clear indications for this follow, for example, from comparison between the average helical repeat of B-DNA, γ, in the

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-1-107-03493-8 - Biophysics of DNA
Alexander Vologodskii
Excerpt
More information

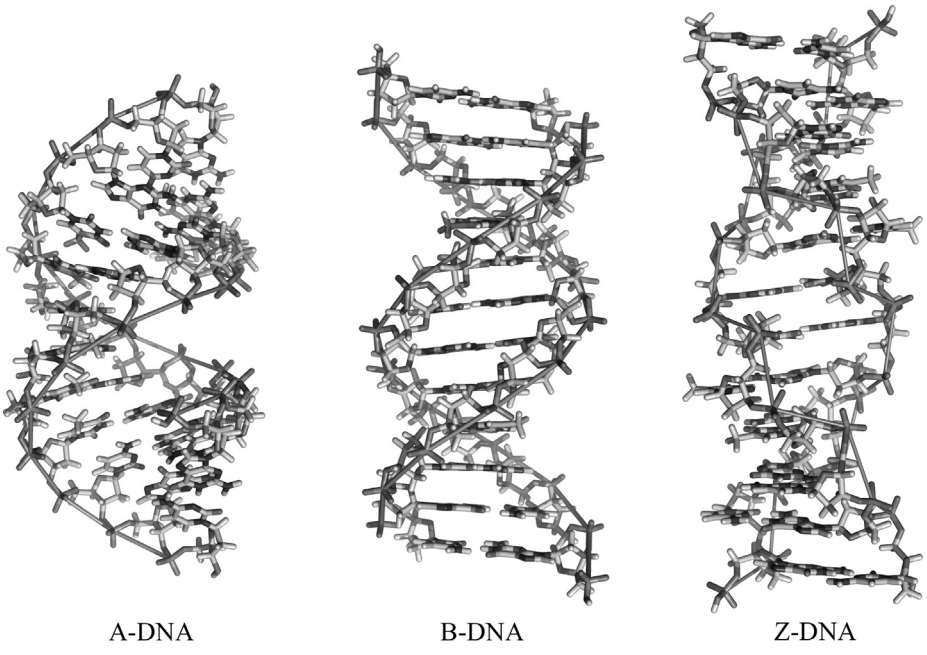

Figure 1.5 Structure of A-, B- and Z-DNA. The phosphorus atoms are connected by a thin orange line to emphasize the geometry of the backbones. Image made by Richard Wheeler, published with permission. A black-and-white version of this figure will appear in some formats. For the color version, please refer to the plate section.

crystal structures and in solutions. While the X-ray data show that the average value of γ is equal to 10.0 bp/turn, it is close to 10.5 bp/turn in solution (see Section 3.2 for details). On the other hand, solution methods have not given sufficient information on the sequence variation of the conformational parameters. In general, thermal fluctuations of the helix parameters exceed their sequence variations (see Section 3.2).

1.2.2 A-DNA

A-DNA is also a right-handed double helix with the same rules of base pairing. B-DNA converts into the A form of the double helix in solutions with low water activity, such as concentrated solutions of ethanol. It is also formed in DNA fibers at humidities of less than 75%. A-DNA has also been found in cocrystals of some DNA–protein complexes. Since RNA duplexes cannot adopt the B conformation, the A form is their normal conformation in solution. A-DNA has 11 bp/turn (Krylov *et al.* 1990, Dickerson & Ng 2001). The base pairs are displaced from the helix axis, and this results in a larger helix diameter, 2.3 nm. The base pairs are tilted by 20° with respect to the helix axis (see Fig. 1.5). In A-DNA they form a propeller with even larger angle than in B-DNA, 19°, between the planes of the bases. It has smaller rise/turn, the distance between base pairs projected to the helix axis, 0.25 nm versus 0.34 nm in B-DNA. The sugar has

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-1-107-03493-8 - Biophysics of DNA
Alexander Vologodskii
Excerpt
More information

$C^{3'}$-*endo* conformation in this DNA form, and the glycosyl angles correspond to *anti* conformations.

1.2.3 Z-DNA

DNA conformation in the Z form is very different from those in the A and B forms (Rich *et al.* 1984). Z-DNA is a left-handed double helix rather than a right-handed one (Fig. 1.5). Its repeating unit consists of two base pairs rather than one. The line connecting the phosphorus atoms does not form a smooth curve, as it does in A- and B-DNA; it zigzags (initiating the name "Z form"). The glycosyl angles in each strand of Z-DNA alternate between *syn* and *anti* conformations, and in each base pair one base is in *anti* and the other is in *syn* conformation. The winding angle (the angle of the helix rotation between adjacent base pairs) also alternates between $-9°$ and $-51°$, for the *anti–syn* and *syn–anti* contacts, respectively. Thus, there are 12 bp/turn in Z-DNA. The sugar conformations also alternate along each polynucleotide chain: $C^{2'}$-*endo* and $C^{3'}$-*endo* correspond to *anti* and *syn* conformations of the nucleotide.

Each base pair has two sides. In Z-DNA the base pairs are inverted relative to the direction of the sugar–phosphate chains, as compared with A- and B-DNA. This inversion changes the conformation of the glycosyl angle for purines. In the case of pyrimidines the sugar is rotated together with the base, so they remain in *anti* conformations. Clearly, the necessity to rotate base pairs makes the transition between B and Z forms more difficult.

A linear double helix can adopt the Z form only at very high concentration of sodium ions, or in the presence of some multivalent ions (Rich *et al.* 1984). Even in these conditions Z-DNA can be formed nearly exclusively in duplexes with alternating purine–pyrimidine sequences. This very strong sequence requirement results from the very high energy of the *syn* conformation of pyrimidines. In DNA segments with alternating purine–pyrimidine sequences, all pyrimidines are in *anti* conformations while all purines are in *syn*. The most favorable sequence for Z-DNA is $d(GC)_n \cdot d(GC)_n$. Still, the sequence requirements are not strict, and under sufficient negative supercoiling many segments of natural DNA adopt the Z form (see Sections 2.4 and 6.5).

1.2.4 B′-DNA

B′-DNA is a conformation that is adopted by poly(dA)poly(dT) in a wide range of solution conditions. Although the structure is close to the B form, it has a narrower minor groove and a large positive propeller twist of the base pairs (Alexeev *et al.* 1987). It has been established that any tract with the sequence $d(A_mT_{n-m})$ adopts this conformation if $n \geq 4$ (see Hagerman (1990) for a review). B′-DNA is able to introduce substantial intrinsic curvature into DNA molecules. The bends appear, mainly, at the stacks between B and B′ helices (Hagerman 1986, Koo *et al.* 1986). For the track of six As in a row the total bend is close to 19° (Koo *et al.* 1990, MacDonald *et al.* 2001). If these bends are repeated in phase with DNA helical periodicity, they can introduce a substantial intrinsic bend in the double helix. This intrinsic curvature was first discovered

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-1-107-03493-8 - Biophysics of DNA
Alexander Vologodskii
Excerpt
More information

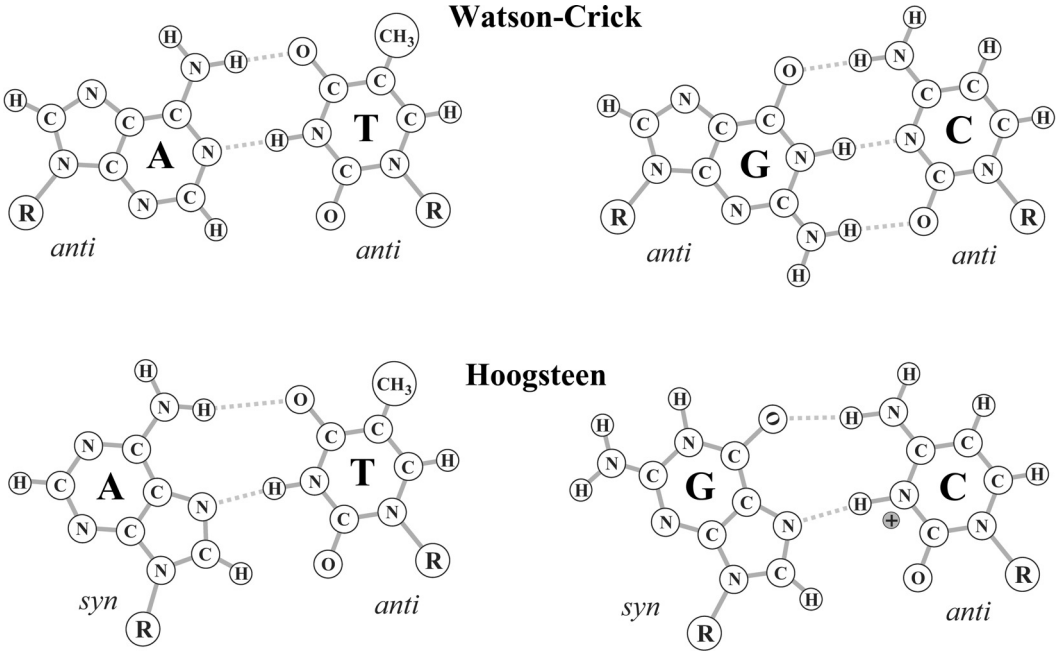

Figure 1.6 Watson–Crick and Hoogsteen base pairing.

in the fragments isolated from kinetoplast DNA that contains phased A-tracts of five to eight nucleotides (Marini *et al.* 1982). It was surprising that X-ray studies of A-tracts did not detect DNA bends observed in solution. The puzzle was solved when it was found that in solutions of 2-methyl-2,4-pentanediol, the dehydrating agent commonly used in crystallization of oligonucleotides, the intrinsic curvature associated with the A-tracts nearly disappears (Sprous *et al.* 1995). Probably this happens due to B′–B transition in the A-tracts.

In addition to the canonical base pairs, many other patterns of base pairing are possible, and some of them have been observed experimentally (Egli & Saenger 1984, Sinden 1994). Some patterns were found in pairs formed by modified bases or in DNA–protein complexes. Probably the most important among these are Hoogsteen and reverse Hoogsteen base pairs (Fig. 1.6), since they were found in triple helices.

1.3 Triple helices

A third polynucleotide chain can be incorporated in the major groove of B-DNA. Although the double helix participating in the triplex formation changes its conformation to some extent, it maintains the major features of B-DNA. In particular, the glycosyl angles correspond to *anti* conformations and the sugars are in C2′-*endo* conformations. The helical repeat of the triple helix corresponds to 12 triads per helix turn (Raghunathan

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-1-107-03493-8 - Biophysics of DNA
Alexander Vologodskii
Excerpt
More information

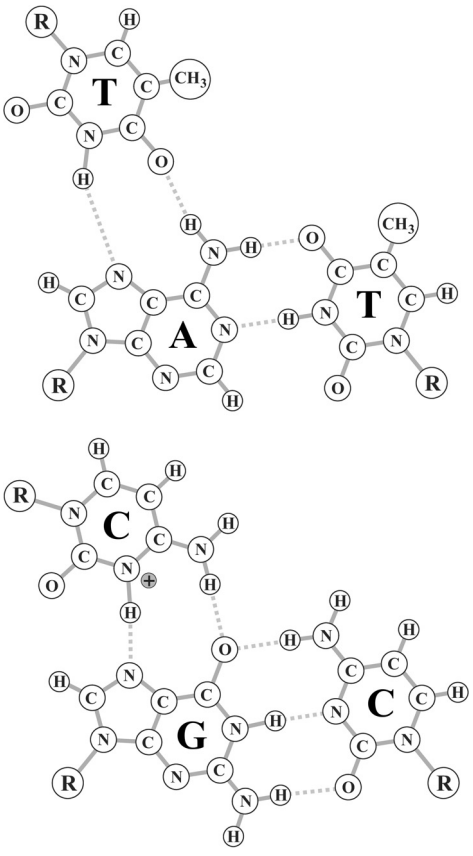

Figure 1.7 Base triads in the Hoogsteen triplexes. The cytosine of the third strand is protonated in the N3 position.

et al. 1993). Over all the varieties of the triple helices the third strand has to form hydrogen bonds with purines of the double helix (Frank-Kamenetskii & Mirkin 1995). This requirement, due to structural restrictions, reduces the types of sequence of B-DNA able to form the triple-stranded helices to homopurine–homopyrimidine segments. The patterns of hydrogen bonding between purines of the B-like DNA and bases of the third strand correspond to Hoogsteen and reverse Hoogsteen base pairs (where the planes of the bases have the same orientations as in Watson–Crick base pairs). According to these patterns, the triple helices can be divided into two groups.

Triplexes with Hoogsteen base pairing. The third strand consists of all pyrimidines and forms Hoogsteen base pairs TA and CG. The two strands connected by Hoogsteen base pairs have to be parallel to each other. Thus, the third strand has to be antiparallel to the pyrimidine strand of the Watson–Crick duplex, and its sequence has to be the reverse of the sequence of that strand. The patterns of hydrogen bonds in the triads of Hoogsteen triplexes are shown in Fig. 1.7. It is important that cytosine participating in the Hoogsteen pairing is protonated in the N3 position.

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-1-107-03493-8 - Biophysics of DNA
Alexander Vologodskii
Excerpt
More information

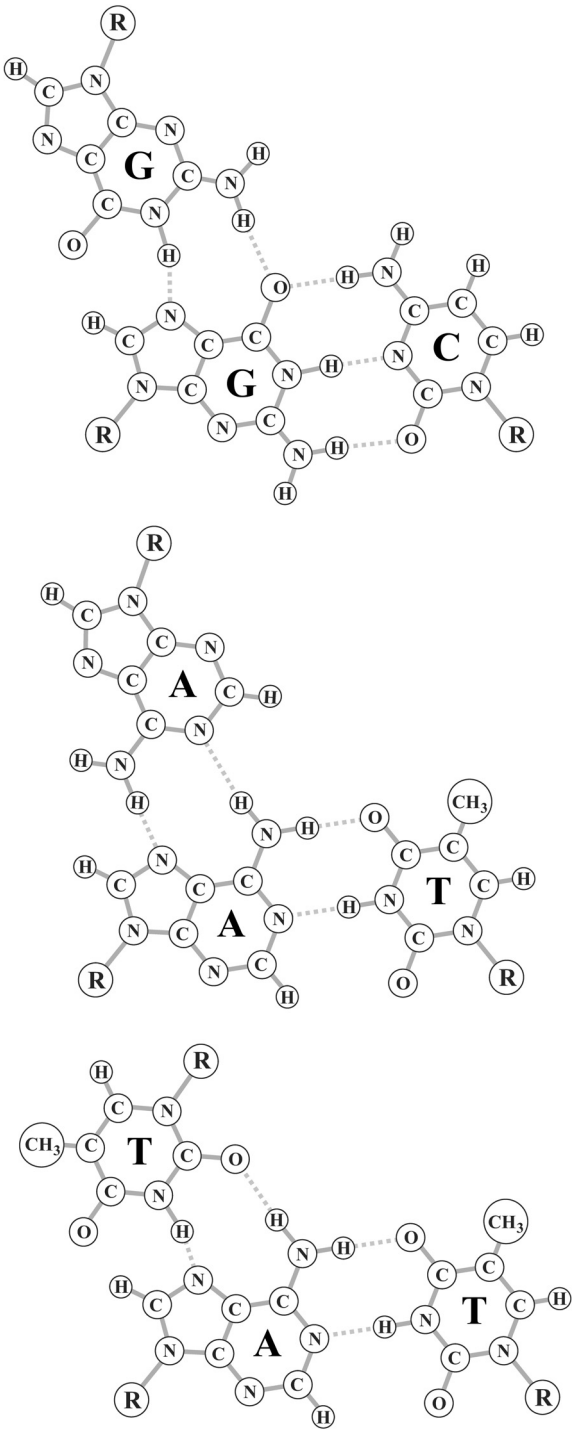

Figure 1.8 Base triads in the reverse Hoogsteen triplexes.

© in this web service Cambridge University Press www.cambridge.org