

Backbone-base inclination as a fundamental determinant of nucleic acid self- and cross-pairing

Pradeep S. Pallan¹, Paolo Lubini², Martin Bolli³ and Martin Egli^{1*}

¹Department of Biochemistry, School of Medicine, Vanderbilt University, Nashville, TN 37232, USA, ²Alta Scuola Pedagogica, CH-6600 Locarno and ³Actelion Pharmaceuticals Ltd., CH-4123 Allschwil, Switzerland

Received July 5, 2007; Revised July 25, 2007; Accepted July 26, 2007

ABSTRACT

The crystal structure of the duplex formed by oligo(2',3'-dideoxy- β -D-glucopyranosyl)nucleotides (homo-DNA) revealed strongly inclined backbone and base-pair axes [Egli,M., Pallan,P.S., Pattanayek,R., Wilds,C.J., Lubini,P., Minasov,G., Dobler,M., Leumann,C.J. and Eschenmoser,A. (2006) Crystal structure of homo-DNA and nature's choice of pentose over hexose in the genetic system. *J. Am. Chem. Soc.*, 128, 10847–10856]. This inclination is easily perceived because homo-DNA exhibits only a modest helical twist. Conversely, the tight coiling of strands conceals that the backbone-base inclinations for A- (DNA and RNA) and B-form (DNA) duplexes differ considerably. We have defined a parameter η_B that corresponds to the local inclination between sugar-phosphate backbone and base plane in nucleic acid strands. Here, we show its biological significance as a predictive measure for the relative strand polarities (antiparallel, *aps*, or parallel, *ps*) in duplexes of DNA, RNA and artificial nucleic acid pairing systems. The potential of formation of *ps* duplexes between complementary 16-mers with eight A and U(T) residues each was investigated with DNA, RNA, 2'-O-methylated RNA, homo-DNA and p-RNA, the *ribopyranosyl* isomer of RNA. The thermodynamic stabilities of the corresponding *aps* duplexes were also measured. As shown previously, DNA is capable of forming both *ps* and *aps* duplexes. However, all other tested systems are unable to form stable *ps* duplexes with reverse Watson–Crick (rWC) base pairs. This observation illustrates the handicap encountered by nucleic acid systems with inclinations η_B that differ significantly from 0° to form a *ps* rWC paired duplex. Accordingly, RNA with a backbone-base inclination of –30°, pairs strictly in an *aps* fashion. On the other hand,

the more or less perpendicular orientation of backbone and bases in DNA allows it to adopt a *ps* rWC paired duplex. In addition to providing a rationalization of relative strand polarity with nucleic acids, the backbone-base inclination parameter is also a determinant of cross-pairing. Thus, systems with strongly deviating η_B angles will not pair with each other. Nucleic acid pairing systems with significant backbone-base inclinations can also be expected to display different stabilities depending on which terminus carries unpaired nucleotides. The negative inclination of RNA is consistent with the higher stability of duplexes with 3'- compared to those with 5'-dangling ends.

INTRODUCTION

Establishing the relative orientation of the helix axis and the dyad that relates the sugar-phosphate portions of the two strands was a key step on the way to a model for the structure of the DNA double helix (1). X-ray diffraction photographs of crystalline sodium thymonucleate fibers indicated a dyad perpendicular to the fiber axis (2) and, accordingly, the two chains of the double helix run in opposite directions. Many years later, DNA duplexes with parallel orientation of strands were proposed (3) and later shown to exist for sequences containing exclusively A and T residues (4,5). The parallel-stranded (*ps*) duplex is destabilized by about 1°C (ΔT_M) per base pair relative to the antiparallel-stranded (*aps*) duplex (6) and is destabilized more strongly by the incorporation of G–C base pairs (7). The hydrogen-bonding structure in *ps* duplex DNA was established as being of the reverse Watson–Crick (rWC) type (8). Although the structure of the *ps* AT-rich double helix has been studied to some extent (9), its biological role remains unknown. Several other structure determinations have provided experimental evidence for double-stranded nucleic acid molecules with parallel orientation of strands and rWC base pairing. These include 5'-CGA containing

*To whom correspondence should be addressed. Tel: +1 615 343 8070; Fax: +1 615 322 7122; Email: martin.egli@vanderbilt.edu

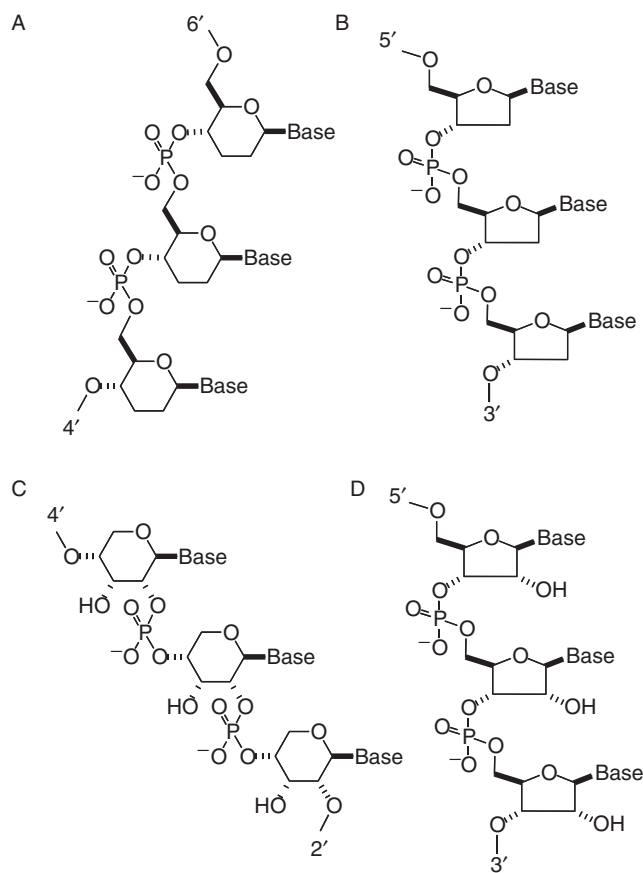


Figure 1. Constitution and configuration of native and artificial nucleic acid pairing systems. (A) β -D-2',3'-dideoxyglucopyranosyl (6' \rightarrow 4')-linked oligonucleotides (homo-DNA), (B) DNA, (C) β -D-ribosepyranosyl (4' \rightarrow 2')-linked oligonucleotides (p-RNA) and (D) RNA.

oligonucleotides (10), a *ps* duplex with isoguanine–cytosine and guanine–isocytosine base pairs (11) and the tetrameric C-rich *i*-motif featuring two self-intercalated *ps* duplexes (12–15). In the present article, we will use the term rWC to refer to base pairs with a *trans* orientation of glycosidic bonds irrespective of whether they are purine–purine, pyrimidine–pyrimidine or pyrimidine–purine pairs.

Experimental investigations of artificial oligonucleotide pairing systems have focused on the pairing properties of a β -D-2',3'-dideoxyglucopyranosyl (6' \rightarrow 4')-linked oligonucleotide analog of DNA (homo-DNA) (16–23), and a β -D-ribosepyranosyl (4' \rightarrow 2')-linked oligonucleotide analog of RNA (p-RNA) (24–26), among others (27) (Figure 1). These studies revealed that homo-DNA constitutes an autonomous pairing system in the sense that homo-DNA oligonucleotides do not pair with any other complementary nucleic acid system. Further, the pairing rules for the four natural bases in the homo-DNA series are not the same as in DNA, suggesting that the WC base pairing rules are a consequence of the base properties as well as of the furanose rings in the DNA backbone. Homo-DNA exhibits WC purine–pyrimidine pairing that is systematically stronger than the corresponding pairings in DNA, and, unlike DNA,

features a reverse-Hoogsteen type purine–purine pairing (18,19). This latter property is not shared by p-RNA that in turn shows stronger adenine–uracil pairing than corresponding RNA duplexes (24). Based on qualitative conformational analyses (16,24) and experimental evidence from solution NMR spectroscopy (20,26), both homo-DNA and p-RNA duplexes were expected to adopt more linear structures, lacking the typical twist observed with DNA and RNA duplexes. The recent crystal structure of a homo-DNA octamer duplex revealed an average twist of $\sim 15^\circ$ for the right-handed duplex (28). Interestingly, the existence of non-helical or N-form duplexes was also suggested for DNA [(29) and literature cited therein]. The energy of such computer-generated and energy-minimized linear dG₁₂–dC₁₂ models exceeded those based on crystallographically determined right- and left-handed duplex forms by 3.5–6.6 kcal/mol per nucleotide. Three of the five backbone torsion angles that were varied (δ angles were kept near a C3'-*endo* conformation, typical for A-form duplexes) were close to those adopted by the GpC step in left-handed Z-DNA. This is in accordance with the low twist value of around -10° displayed by purine–pyrimidine steps in Z-DNA. The most notable difference from the standard A- and B-DNA backbone torsion angles in the linear models was found for angle β which is found in either a *synclinal* + or a *synclinal* – conformation, but is strictly *antiperiplanar* in all crystal structures of A- and B-DNA duplexes and drug–DNA complexes. It was speculated that, although less stable than the helical forms of DNA, such linear DNA structures might be associated with regulatory regions of active genes (29).

DNA structures are commonly characterized with several geometrical parameters, such as helical rise and twist, inclination, displacement, roll, propeller twist, sugar pucker, backbone and glycosidic torsion angles, groove widths, helical bend, etc. (30). Thus, A-DNA, the low-humidity form, displays the typical tilting of base pairs as a result of the contraction of the duplex along the helix axis at lower levels of hydration, while the base pairs in the B-DNA duplex are oriented more or less normal to the helix axis. Typically, the extracted rotational and translational parameters involve the two bases of a base pair (e.g. propeller twist and shear) or two successive base pairs (e.g. twist and rise) (31). In any case, the relative orientations of base pairs depend on the choice of a local or global helix axis. In general, the defined parameters allow one to classify a right-handed double-helical nucleic acid fragment as either an A-, or a B-type. Occasionally, duplexes may reveal geometrical features that are distinct from either canonical A-DNA or B-DNA. Thus, a particular DNA conformation with an enlarged major groove, found in a series of crystal structures of protein–DNA complexes, has prompted the classification of duplexes based on the spatial relationship of the base pairs, the phosphate backbone and the helix axis (32). Unlike in the native right-handed duplexes with helical intertwined strands, the relative orientations of backbone and bases can be readily discerned in linear (20,29,33) or strongly unwound duplexes (28,34). The considerable

backbone-base inclination in structural models of non-helical duplex DNA has been noted (29,35–37), but to the best of our knowledge, nobody has ever defined a parameter that would provide a measure for the backbone-base inclination in either the canonical *aps* A-DNA/A-RNA and B-DNA duplexes or model duplexes in general.

The orientation of a base pair relative to either a local or overall helix axis is described by the inclination angle η (referred to as tilt angle earlier on) (31). In duplexes of the A-type, η is usually between 10 and 20° and it is around 0° for B-type duplexes (38). With reference to this base pair-helix axis inclination angle η , we have defined an angle η_B (backbone) that measures the inclination between the sugar-phosphate backbone and single bases or base pairs in a double-stranded oligonucleotide. Evidently, the corresponding inclinations in single- or multiple-stranded arrangements can also be assessed. We have calculated the backbone-base pair inclinations in various duplex types formed by DNA, RNA, homo-DNA and p-RNA, as well as in triple- and four-stranded motifs. This analysis reveals that A- and B-form duplexes display fundamentally different degrees of backbone-base inclinations. Based on UV-melting experiments, we have found that while DNAs containing only A and T bases form stable *ps* rWC base-paired duplexes, RNA, homo-DNA and p-RNA do not. The underlying cause for the distinct tendencies of DNA and RNA to adopt *ps* duplexes with rWC pairing resides in their different backbone-base inclinations. The loss of hydrogen-bonding stabilization for a general sequence containing all four bases when switching from a WC to an rWC pairing mode (3 versus 2 hydrogen bonds, respectively, in a G–C base pair) and the limited conformational flexibility yielded by the *ps* duplex (which is only accessible when the bases are oriented in a more or less normal way to the backbone) are the chemical physical origins of the preference for an *aps* orientation of strands in the DNA double helix. Conversely, in the case of RNA and the more linear pyranose-based oligonucleotide systems, the formation of *ps* rWC paired duplexes can be excluded solely on the grounds of their intrinsic strong backbone-base inclinations. Here, we describe the computation of the η_B inclination angle for various pairing motifs formed by native and artificial nucleic acid systems and highlight its importance as a fundamental determinant of nucleic acids self- and cross-pairing.

MATERIALS AND METHODS

Calculation of the backbone-base inclination angle η_B

For the computation of the inclination angle η_B between a base pair and the backbone of a nucleic acid system, the following algorithm was used (see Figure 2). First a BSpline curve (red) based on the coordinates of the nucleotide phosphorus atoms (orange dots) was calculated for each single strand independently (39). The positive direction was chosen as the one pointing from the 5'-terminus toward the 3'-terminus. Next, the

best plane α through the atoms of a particular base pair (white) as well as the pair's center of mass was calculated and the point where the BSpline curve penetrates α was determined (blue dot on red curve). This allows the definition of a second plane β (cyan) that is perpendicular to α , contains both points calculated above, the center of gravity and the point of penetration, and of course the normal to α through the center of gravity of the base pair (brown vector, pointing in the direction of the 3'-terminus). The inclination angle η_B is defined as the angle between the vector normal to α and the projection (dotted line in yellow) on β of the tangent (green vector, pointing in the direction of the 3'-terminus) to the BSpline curve at the point where the latter penetrates α . η_B is positive if the scalar product between the projected tangent vector (defined by the two blue dots in a 5'→3' direction) and the vector from the BSpline/ α point of penetration to the center of mass is negative. In Figure 2 (left panels), η_B is equal to the angle between the brown (normal) vector and the blue (local backbone direction) vector and therefore corresponds to ca. –40° for A-DNA and ca. 0° for B-DNA.

Fifty-four different DNA duplexes with lengths between 8 and 12 base pairs with standard Watson–Crick base pairing, whose structures were determined by single crystal X-ray diffraction methods to resolutions between 1.3 and 2.8 Å, were retrieved from the NDB database [(40), <http://ndbserver.rutgers.edu>] and the inclination angles η_B for individual base pairs were calculated following the procedure described above. The η_B angles of terminal base pairs were not included in the analysis as their geometries are often affected by considerable conformational freedom. The following A- and B-DNA crystal structures (NDB entry codes) were used to generate the histograms depicted in Figure 3 (duplexes adopting exact 2-fold crystallographic symmetry are designated with an asterisk; DNA–RNA chimeric A-form duplexes are designated with a superscript c): ADH006, ADH007, ADH008*, ADH010, ADH012*, ADH014*, ADH018, ADH019, ADH020*, ADH023*, ADH027*, ADH031, ADH038*, ADH054, ADH059, ADI009, ADJ022*, ADJ049, ADJ050, ADJ051*, ADL025, ADL045*, ADL046*, AHJ015^c, AHJ040^c, AHJ043^c, AHJ044^c, AHJ052^c, AHJS55^c, BDJ008*, BDJ025, BDJ031, BDJ036, BDJ039, BDJ051, BDJ052, BDJ055, BDJB43, BDJB44*, BDJB48, BDJB50, BDL002, BDL006, BDL007, BDL009, BDL011, BDL012, BDL014, BDL015, BDL021, BDL022, BDL028, BDL042 and BDL047.

Inclination angles listed in Table 1 for two-, three- and four-stranded nucleic acid molecules as well as artificial pairing systems were also calculated using the BSpline approach.

Oligonucleotide syntheses and purification

DNA and 2'-O-methylated RNA 16-mer oligonucleotides were purchased from Microsynth Ltd., Balgach, Switzerland. The corresponding homo-DNA, pRNA and RNA oligonucleotides were synthesized according to described procedures (17,24). Fully deprotected oligonucleotides were purified with RP-HPLC

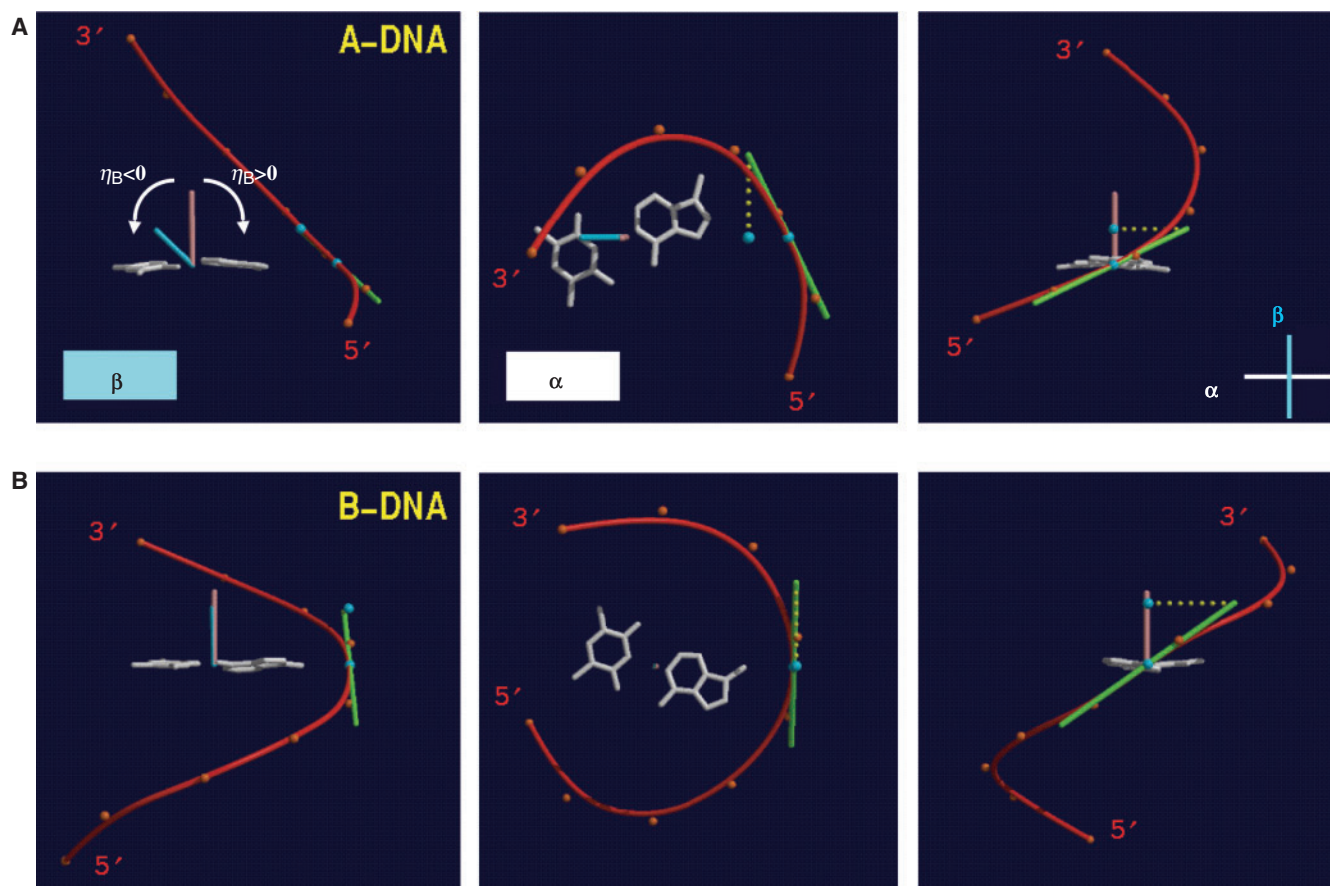


Figure 2. Definition of the backbone-base pair inclination angle. The relative orientations of a base pair (A–T, white) and the BSpline backbone curve (red) as defined by the phosphorus atoms (orange dots) for (A) A-DNA and (B) B-DNA. The three images each depict projections along the base pair and roughly normal to the WC hydrogen bonds (left), along the normal (brown) to the best plane through the base pair (center), and roughly along the long axis of the base pair (right). For further details see the description of the calculation of the backbone-base inclination angle η_B in the experimental procedures.

(buffer A: 100 mM Et_3NHOAc in H_2O , buffer B: 100 mM Et_3NHOAc in mixtures of $\text{H}_2\text{O}/\text{CH}_3\text{CN}$ of various ratios). Following desalting on Sep-Pak columns, oligonucleotide-containing fractions were combined and evaporated to dryness. All 16-mers were characterized by analytical HPLC and MALDI-TOF MS and were $\geq 95\%$ pure.

UV MELTING TEMPERATURE (T_M) MEASUREMENTS

The concentrations of 1:1 mixtures of strands were adjusted to ca. $1.0\ \mu\text{M}$ (duplex) and the buffer was 150 mM NaCl, 10 mM Tris–HCl (pH 7.0). UV-melting curves for mixtures and single strands alone were recorded on a Perkin Elmer Lambda 2 Spectrometer, equipped with a Perkin Elmer Digital Controller/Temperature Programmer C570. The temperature was varied between 4°C and 95°C (measured directly in the sample solution) with a temperature gradient of $\sim 0.8^\circ\text{C}/\text{min}$. The hyperchromicities and hypochromicities were measured at 250, 260, 270 and 280 nm during two heating cycles and one cooling cycle, respectively. All curves were analyzed with the program Kaleidagraph (version 3.0 for MacIntosh).

RESULTS AND DISCUSSION

Backbone-base pair inclination in standard nucleic acid duplexes

Analogous to the calculation of the inclination angle η between a base or base-pair plane and the helix axis, where the actual value depends on the definition of the latter (31), inclination angles η_B depend on how the course of the sugar-phosphate backbone is being traced. A simple way to follow a nucleic acid backbone is by connecting successive phosphorus atoms along a strand. Similarly, the $\text{C1}'$ sugar atoms can serve as reference points. A somewhat more sophisticated approach is based on the calculation of a so-called BSpline curve (39) that smoothly traces either the phosphorus or the $1'$ -carbon positions. The BSpline approach using P atoms is illustrated in Figure 2. Details concerning the computation of the backbone-base pair inclination angle η_B are given in the Experimental section.

We selected 54 crystal structures of A- and B-DNA duplexes from the Nucleic Acid Database (40) (see Experimental section for entry codes) and calculated the η_B angles of individual base pairs for both strands using all four of the above methods to define the

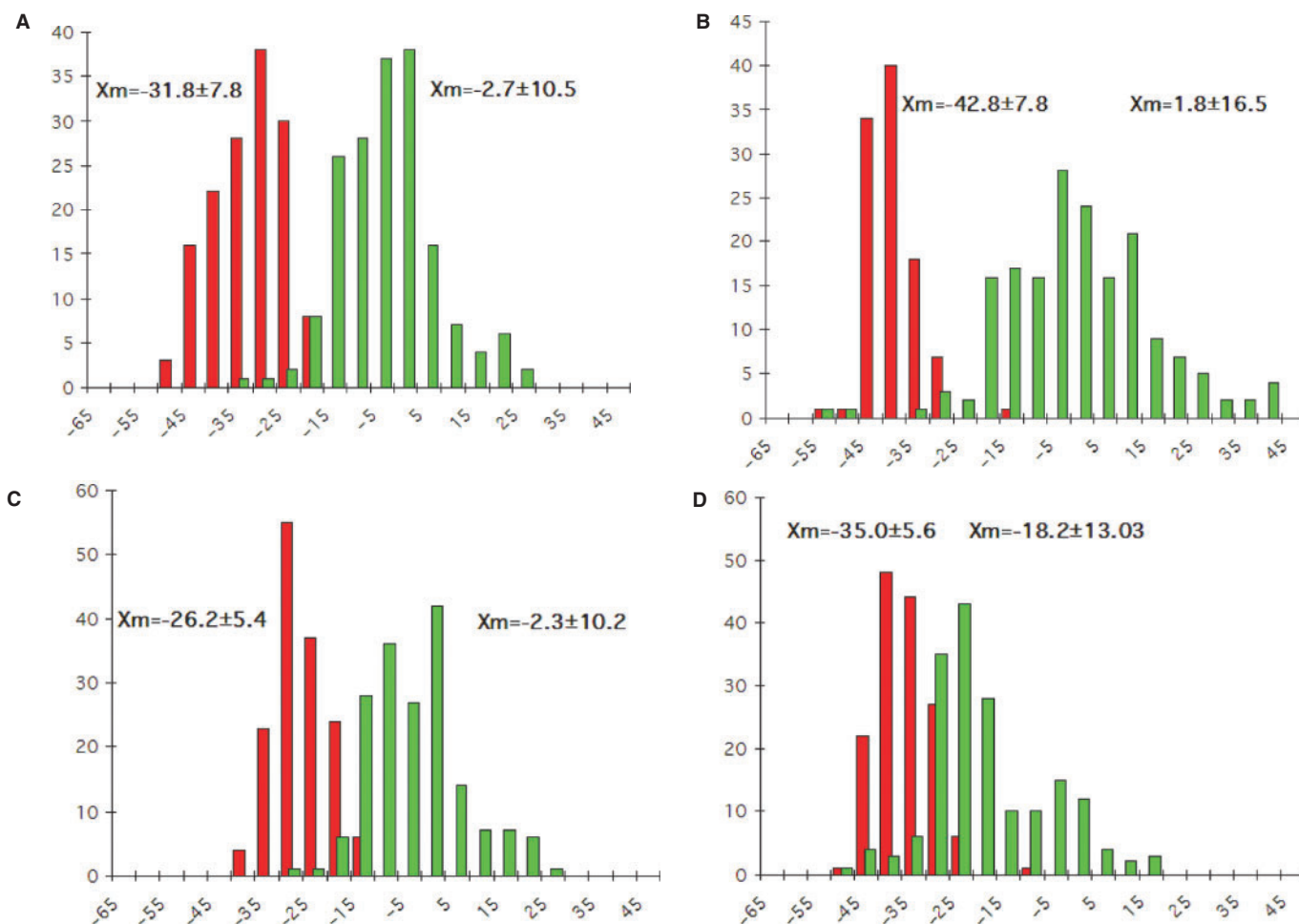


Figure 3. Backbone-base pair inclinations in A-DNA (red) and B-DNA (green) duplexes. The definition of the backbone direction affects the backbone-base pair inclination angle η_B : (A) calculating a BSpline curve through P atoms; (B) vectors connecting P atoms from adjacent residues along the strand; (C) calculating a BSpline curve through C1' atoms and (D) vectors connecting C1' atoms from adjacent residues along the strand. The average inclinations and SDs for A- and B-form duplexes are given above the histograms. The tighter distribution of inclination angles with A-form duplexes is apparent in all panels and is not a consequence of the different numbers of observations included in the analysis (29 A-DNA structures and 25 B-DNA structures; see the Materials and Methods section), but is likely a manifestation of the more limited conformational flexibility of A-DNA relative to B-DNA. Although the graphs represent a compilation of individual base pairs in many duplex structures, the average backbone-base inclination angle based on a single structure is typically of sufficient predictive value for a particular duplex family or class of nucleic acids.

direction of the backbone (P BSpline, $P_n \rightarrow P_{n+1}$ vectors, C1' BSpline, $C1'_n \rightarrow C1'_{n+1}$ vectors). Histograms summarizing the results of the analysis are depicted in Figure 3. Apart from the results based on C1' \rightarrow C1' vectors for defining the backbone (Figure 3D), the histograms nicely demonstrate the different backbone-base inclinations in A-DNA and B-DNA (see also Figure 2). Almost complete separation of the angle regions covered by the two duplex types is observed when defining the backbone simply as a series of straight lines connecting the phosphorous atoms of a strand (Figure 3B). However, the angle values based on the other backbone definition methods reveal similar tendencies concerning the inclination in the duplexes. In the further course of the discussion, we will thus only consider the angle values calculated using a BSpline curve through the phosphorus atoms.

The analysis reveals a clear difference between the backbone-base pair inclinations in A-DNA and B-DNA

duplexes. Although B-DNA duplexes display a rather wide range of inclinations, the majority of base pairs in this duplex type exhibit inclinations between -15° and $+15^\circ$ with a mean value near 0° . Not only are the base pairs roughly perpendicular to the helix axis in B-type duplexes, but the same consequently holds true also for the relative orientation between base pairs and backbone. In A-DNA duplexes, the inclination angle distribution is somewhat narrower compared with the B-form, with an average of around -32° . This is consistent with the more uniform character of A-DNA and the general geometrical and topological similarities between the crystal structures of A-DNA duplexes [(41) and references cited therein]. Base pairs in A-DNA are therefore inclined with regard to both, the backbone and the helix axis. However, the two inclination types differ significantly for A-form DNAs and therefore also for RNA duplexes: inclination angles η between base pairs and helical axis are positive and

Table 1. Average backbone-base pair inclinations η_B in non-standard nucleic acid structures, involving parallel (*ps*) and antiparallel (*aps*) orientation of strands and various base pairing modes¹

Sequence and synopsis	NDB code	Reference	Base pair	Pairing mode	Strand polarity	η_B (σ) Strand 1	η_B (σ) Strand 2
d(CCCC) self-intercalated i-motif paired via C C ⁺ base pairs	UDD024	(13)	C1–C5 to C4–C8	rWC	<i>ps</i>	+0.9° (4.1°)	+1.8° (2.0°)
d(CGATCG) <i>ps</i> duplex at low pH with homo base pairing	supporting information (NMR solution structure)	(10)	C9–C13 to C12–C16	rWC	<i>ps</i>	+2.7° (0.1°)	+0.6° (2.5°)
			G2–G8 to C5–C11 (PSte model ²)	rWC	<i>ps</i>	–1.5° (12.8°)	–1.5° (12.8°)
d(CGCGCG) ³ left-handed Z-DNA duplex	ZDF035	(45)	G2–G8 to C5–C11 (PSts model ²)	rWC	<i>ps</i>	–0.4° (13.1°)	–0.4° (13.1°)
d(ATATAT) duplex with Hoogsteen pairs only	UD0035	(46)	T2–A11 to A5–T8	Hoogsteen	<i>aps</i>	–7.3° (6.0°)	–6.0° (5.1°)
r(GACACCUGAUUC)/d(GAATCAGGTGTC) RNA–DNA hybrid bound to RNase H	PH0016	(47)	rA2–dT11 to rU10–dA3	WC	<i>aps</i>	–25.2° (5.6°)	–33.2° (9.0°)
dd(CGAATTCG) homo-DNA duplex	UD0070	(28)	G2–C15 to C7–G10	WC	<i>aps</i>	+34.3° (10.9°)	+35.5° (10.8°)
pr(CGAATTCG) pyranosyl-RNA duplex	– (NMR solution structure)	(26)	G2–C15 to C7–G10	WC	<i>aps</i>	–46.2° (3.8°)	–46.2° (4.2°)
d(TGGGGT) parallel-stranded tetraplex with four G-quartets	UDF036	(48)	G2–G12 to G5–G15	Hoogsteen (adjacent strands)	<i>ps</i>	+19.4° (11.7°)	–20.6° (3.0°)
			G2–G32 to G5–G35	Hoogsteen (diagonally related strands)	<i>ps</i>	–31.6° (9.5°)	+26.1° (11.6°)
d(GGGG ^{Br} UTTTGGGG) antiparallel-stranded tetraplex with four G-quartets	UD0013	(49)	G1001–G2009 to G1004–G2012	Hoogsteen (adjacent strands)	<i>aps</i>	+17.0° (8.1°)	+29.7° (7.2°)
			G1001–G1012 to G1004–G1009	Hoogsteen (diagonally related strands)	<i>aps</i>	+0.5° (8.4°)	+2.2° (7.4°)
p(CGACG) ⁴ peptide nucleic acid duplex	UPNA56	(50)	G2–C12 to C5–G9	WC	<i>aps</i>	–35.5° (1.6°)	–32.8° (1.5°)
d(TCCTCCTTTTITAGGA GGATTTTITGGTGGT) intramol. Pu–Pu–Py type triplex with G–G–C and T–A–T triples (T loops underlnd.)	Protein Data Bank entry 134D	(52)	C2–G13 to C6–G9	WC	<i>aps</i>	–25.5° (8.9°)	–22.6° (4.4°)
	(NMR solution structure)		G9–G20 to G13–G16	Hoogsteen and reverse Hoogsteen	<i>aps</i> (Pu) <i>ps</i> (Py)	+10.6° (12.5°)	+24.1° (9.3°)

¹Using a Bspline curve for the phosphate backbone.²Parallel Stranded extended model [PSte; rWC like, symmetrical type I (A), type III (G) homo-purine and type XII (T), type XIV (C) homo-pyrimidine pairing] and Parallel Stranded short model [PSts; symmetrical type II (A), type IV (G) homo-purine and type XIII (T), type XV (C) homo-pyrimidine pairing] (30).³Defining the backbone by P → P vectors may be superior to the Bspline approach due to the zigzag shape of the Z-backbone.⁴Right-handed conformation of the PNA duplex; Bspline curve through C5' atoms.

can amount to 20° ; inclination angles η_B between base pairs and backbone are negative and can reach -50° (Figure 3A). These differences are a consequence of the underlying topological properties of a duplex as illustrated by similar values for η and η_B inclinations in B-DNA in contrast to A-DNA, where helix axis and backbone directions diverge to a significant degree.

It also appears that η_B is sensitive to conformational variations within the A-form duplex family. Thus, the average inclination in the crystal structure of an A'-DNA with sequence d(CCCCGGGG) [(42), NDB code ADH012 (40)] is $-35.4^\circ \pm 6.6^\circ$, whereas it is $-40.2^\circ \pm 6.3^\circ$ and $-40.9^\circ \pm 3.4^\circ$ (two numbers in cases where the strands are not related by a crystallographic dyad) in the rhombohedral crystal form of the A-RNA with identical sequence r(CCCCGGGG) [(43), ARH064]. These two structures differ considerably in their geometries. The bases are notably less inclined relative to the helix axis in the DNA duplex, its average rise is 0.6 \AA larger than in the RNA, and its major groove is twice as wide as the one of the RNA (43). However, other A-DNA duplexes have quite similar base pair-backbone inclinations compared with d(CCCCGGGG). For example, the average values for η_B are $-34.1^\circ \pm 7.9^\circ$ and $-34.2^\circ \pm 7.2^\circ$ in strands 1 and 2, respectively, in the crystal structure of a typical A-DNA with sequence d(CCCCGCGGGG) [(44), ADL025]. Also, a variety of other A-RNA duplexes (AHIB53, ARL048, ARN035) show η_B inclinations with average values of ca. -30° , closer to those of the above A- and A'-DNAs than to the one found for the r(CCCCGGGG) A-RNA.

Homo-DNA and p-RNA have strongly inclined backbones

Three-dimensional structures determined for self-complementary homo-DNA [X-ray crystallography (28)] and p-RNA [solution NMR (26)] oligonucleotides with the sequence CGAATTCG revealed duplexes with strongly inclined backbones in both cases. However, the inclinations are of the opposite sign (Figure 4, Table 1), positive for homo-DNA and negative for p-RNA. Homo-DNA adopts a right-handed conformation with an average twist angle between adjacent base pairs of ca. 15° , although the individual twist angles vary greatly. For example, base pairs at the central ApT step show virtually no twisting (Figure 4A). The average backbone inclination angle is about $+35^\circ$. By contrast, the average backbone inclination in p-RNA is -46° . The p-RNA duplex is quasi-linear with a slight left-handed twist. Although the relative strand orientation in both duplexes is antiparallel and the bases are paired in the WC mode, the prevalence of inter-strand stacking that is a direct consequence of the strongly inclined backbones distinguishes these hexopyranosyl duplexes from those formed by DNA and RNA. Whereas homo-DNA was shown to engage in purine-purine self-pairing via the reverse-Hoogsteen mode (19) (consistent with the ability of the 2',3'-dideoxy- β -D-glucopyranosyl-phosphate backbone to twist), the absence of self-pairing involving the Hoogsteen or reverse-Hoogsteen modes in the p-RNA series is

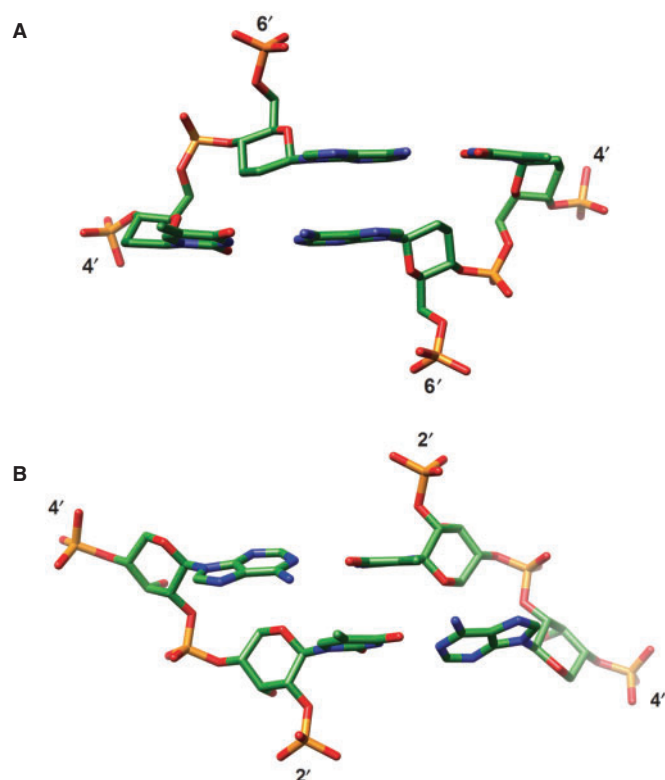


Figure 4. The strictly antiparallel pairing mode of homo-DNA and p-RNA is the result of highly inclined backbone and base-pair axes. The central ApT base-pair steps in the structures of the (A) homo-DNA (28) and (B) p-RNA (26) octamer duplexes with sequence CGAATTCG. The views are into the major groove and illustrate the strong, positive and negative backbone-base pair inclinations in homo-DNA and p-RNA, respectively, as well as the dominance of cross-strand stacking. In both cases, the quasi-linear local geometry of the backbone renders the directions of backbone and helix axis nearly parallel.

noteworthy (25). This observation indicates that p-RNA is more linear and more rigid than homo-DNA. Interestingly, in view of the strong negative backbone-base inclinations exhibited by both RNA (A-form, Figure 3) and p-RNA (Figure 4B), there is no cross-pairing between the two (27).

Backbone-base inclination in alternative DNA structures and PNA

Backbone-base pair inclination angles for alternative structures adopted by DNA or RNA and for a variety of nucleic acid analogs are listed in Table 1. From the two examples of *ps* duplexes (the i-motif adopted by C-rich strands is composed of two interspersed *ps* duplexes) with rWC pairing that were analyzed, we can conclude that molecular dyad and helix axis coincide in such structures. This makes a drastic deviation from a perpendicular orientation between backbone and base pairs impossible. Every double-helical arrangement with a dyad along the helical axis is inevitably more restricted conformationally compared with *aps* duplexes in which local dyads are normal to the helix axis

and positioned within and between base-pair planes. Hence, *aps* DNA duplexes display a range of inclinations (Figure 3) and the left-handed Z-DNA (45) differs from both right-handed duplex families also because of its strongly positive values for η_B (Table 1). Further, the data presented reveal that an *aps* DNA duplex with Hoogsteen-type pairing resembles B-form DNA not only in terms of diameter and groove topology (46) but it also exhibits a similar backbone inclination.

The DNA and RNA strands in a DNA–RNA hybrid duplex bound to RNase H have distinct sugar conformations [Southern, i.e. C2'-*endo*, and Northern, i.e. C3'-*endo*, respectively (47)]. Nevertheless, the duplex is best not thought of as composed of a B-DNA strand paired to an A-RNA strand as the inclination angles of the two are rather similar. Thus, the DNA strand appears to adapt to the conformational constraints of the RNA complement, and apparently does so by retaining sugar puckers that are typically associated with B-DNA. Analysis of the sugar puckers in a DNA strand alone is therefore not a sufficient means to classify its conformation in terms of the A- and B-form types. On the other hand, the backbone-base inclination angle may provide a more meaningful conformational gauge.

Formation of three- and four-stranded DNA motifs with *ps* orientation of strands and Hoogsteen-type base pairing does not limit the inclination angle range to a similar extent as found above for the duplexes with rWC pairing. The backbone-base pair inclinations in the *ps* G-rich tetraplex (48) vary between -32° and $+26^\circ$ (Table 1). Moreover, the *ps*- and *aps*-stranded G-tetraplexes (49) exhibit different inclination angle patterns and in the latter, Hoogsteen-paired strands have inclinations that are strongly positive (17° – 30° , Table 1).

In the so-called P-helix formed by double-stranded peptide nucleic acids [PNA (50)], backbone and base pairs form an angle of ca. -35° . To calculate the inclination angle in PNA, we treated the C5' carbons as phosphorus atoms. Despite a slower writhe compared with RNA (PNA can be either right- or left-handed), PNA's inclination is thus similar to that of duplex-RNA. PNA forms stable *aps* duplexes with both RNA and DNA under formation of WC base pairs. This can be taken as an indication that pairing of two strands (that are either of the same chemical type or represent different chemistries) requires their backbone-base inclinations to be similar. It should be noted that backbone-base inclination alone is insufficient to predict cross-pairing

and that helicality (degree of twisting and helix sense) also need to be taken into account. Thus, homo-DNA, Z-DNA and (left-handed) PNA all exhibit positive backbone inclinations, but the right-handed homo-DNA strand cannot pair with either left-handed DNA or PNA. In DNA–PNA hybrids (51) and in the above DNA–RNA hybrids, the DNA strand adapts to the geometry of the partner strand and not vice versa. DNA's conformational versatility is further demonstrated by the fact that a model of the homo-DNA crystal structure composed entirely of 2'-deoxyribonucleotides could be refined to an R-factor of 34% (28).

Probing the existence of *ps* rWC paired RNA, p-RNA and Homo-DNA duplexes

DNA forms stable *ps* duplexes with rWC pairing (4–8). In order to examine whether RNA (and 2'-*O*-methylated RNA, 2'-OMe-RNA), homo-DNA and p-RNA also exist in this form, we conducted a series of UV-melting experiments, assessing the stabilities of both the standard *aps* duplexes between a 16-mer sense strand and its antiparallel complement, as well as those of the putative duplexes between the same sense strand and its parallel complement. The sequence of the sense strand 5'-TTTAAATATAATAAT (U replaces T in the RNA and 2'-OMe-RNA strands)—was designed so as to avoid self-aggregation, hairpin formation or duplex formation through slippage as much as possible. We did not incorporate any G or C bases in order to guarantee maximum stability of possible *ps* rWC base-paired duplexes. Methylation of the 2'-OH group of RNA results in a considerable stabilization of standard *aps* DNA–RNA (53, 54) and RNA duplexes (55) (ca. $1^\circ\text{C}/\text{per}$ substitution). We speculated that an eventual *ps* duplex formation would also be more favorable for 2'-OMe-RNA relative to RNA. Therefore, we used the 2'-OMe-RNA oligonucleotides in order not to potentially overlook the existence of *ps* rWC paired RNA due to the low stability of the duplex. This is based on the assumption that RNA and 2'-OMe-RNA are alike with regard to their pairing properties and propensities for formation of alternative (*ps*, non-WC base-paired) double-helical structures. Individual sequences and (where measurable) T_m values of the investigated double-stranded arrangements with equal and opposite strand polarities are listed in Table 2. Corresponding UV-melting curves for DNA and RNA are depicted in Figure 5, and Figure 6 shows the experimental data for homo-DNA and p-RNA.

Table 2. T_m values in $^\circ\text{C}$ for duplexes with antiparallel and parallel orientations of strands

System	Linkage	x- TTT TAA ATA TAA TAA T - y y- AAA ATT TAT ATT ATT A - x	x- TTT TAA ATA TAA TAA T -y x- AAA ATT TAT ATT ATT A -y
DNA	x = 5' → y = 3'	28.5	16.5
homo-DNA	x = 6' → y = 4'	67.0	—
pRNA	x = 4' → y = 2'	60.0	—
		x- UUU UAA AUA UAA UAA U - y y- AAA AUU UAU AUU AUU A - x	x- UUU UAA AUA UAA UAA U -y x- AAA AUU UAU AUU AUU A -y
RNA	x = 5' → y = 3'	28.4	—
2'-OMe-RNA	x = 5' → y = 3'	39.8	—

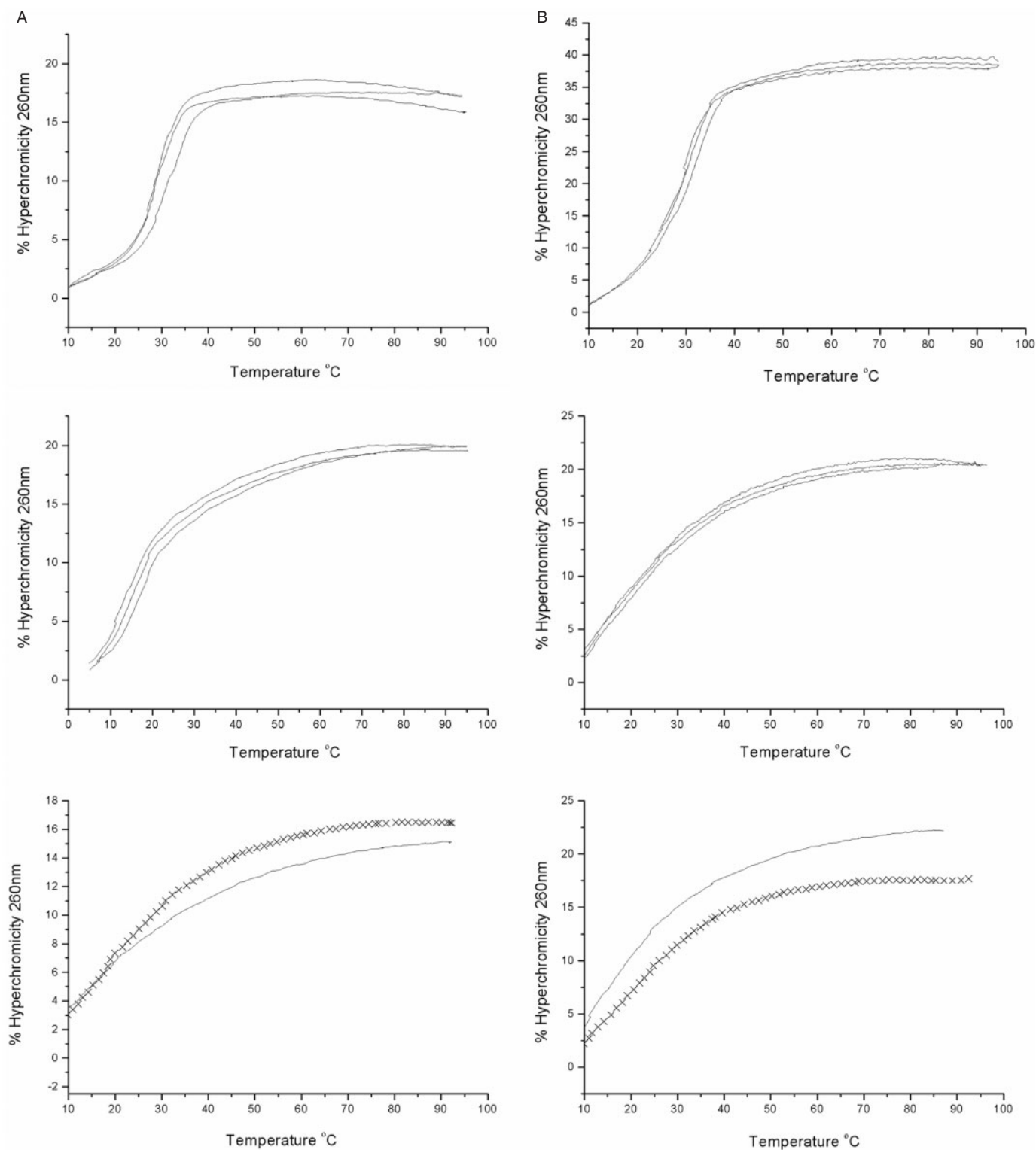


Figure 5. Thermodynamic stabilities of DNA and RNA duplexes with antiparallel and parallel orientation of strands. UV-melting curves for 1:1 mixtures of (A) DNA and (B) RNA 16-mers with opposite (top) and equal (center) strand polarities. Bottom panels show melting profiles for single strands that constitute the parallel-stranded arrangements: 5'-TTT TAA ATA TAA TAA T-3' (DNA, solid lines); 5'-AAA ATT TAT ATT ATT A-3' (DNA, crosses); 5'-r(UUU UAA AUA UAA UAA U)-3' (RNA, solid lines); 5'-r(AAA AUU UAU AUU AUU A)-3' (RNA, crosses). The apparent melting profile of the RNA duplex with *ps*-orientation of strands is the result of the temperature-dependent hyperchromicities of the constituting single strands. For the corresponding UV-melting diagrams with 2'-OMe-RNA, see the Supplementary Data.

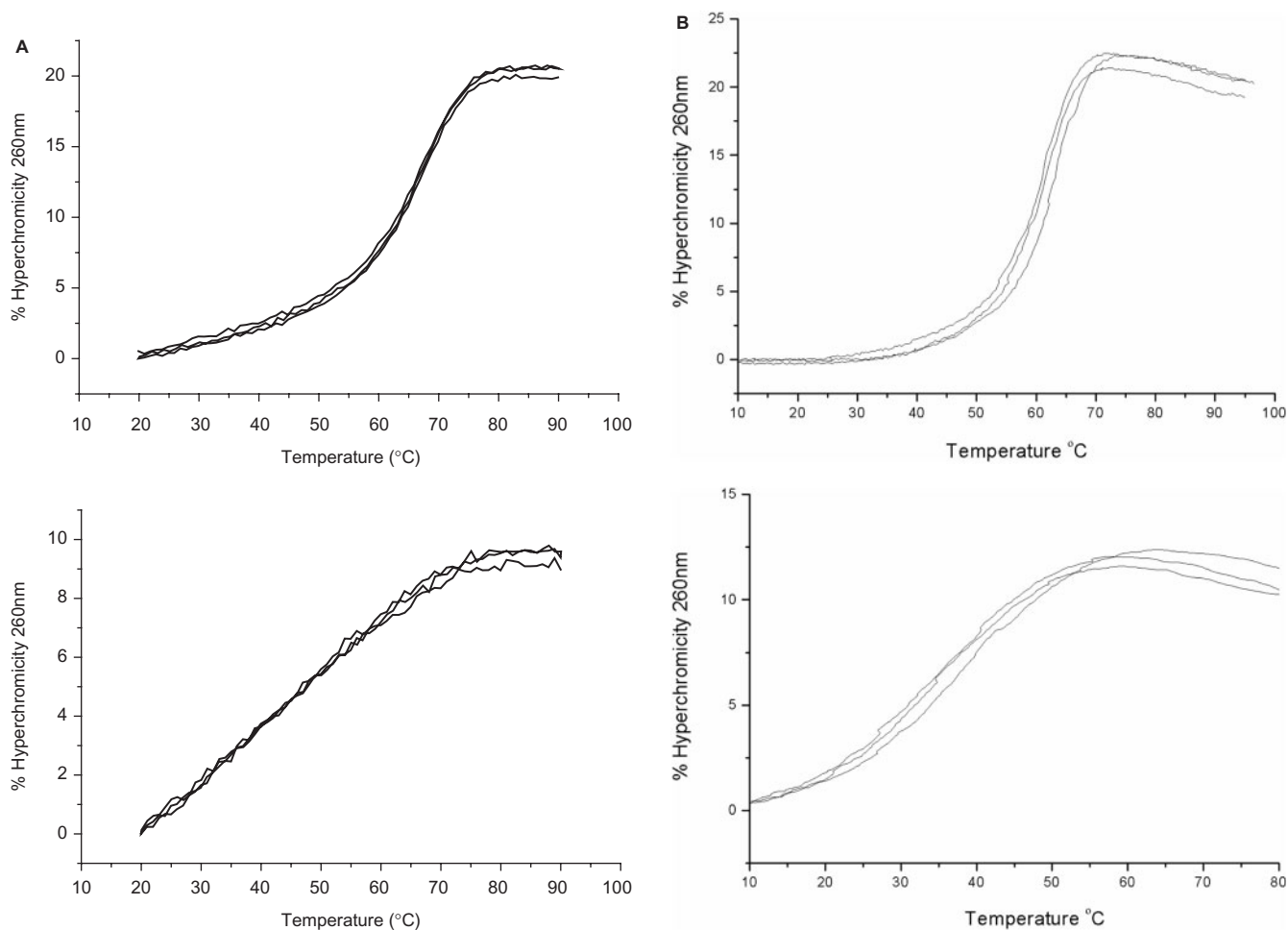


Figure 6. Thermodynamic stabilities of homo-DNA and p-RNA duplexes with antiparallel and parallel orientation of strands. UV-melting curves for 1:1 mixtures of (A) homo-DNA and (B) p-RNA 16-mers with opposite (top) and equal (bottom) strand polarities. The apparent melting profiles of the duplexes with *ps*-orientation of strands are due to the temperature-dependent hyperchromicities of the constituting single strands (data not shown).

In each case, the temperature-dependent hyperchromicities were also measured for the three single strands separately (shown in Figure 5 for DNA and RNA). The recorded data did not indicate any anomalous behavior as far as secondary structure formation is concerned.

From the combined UV-melting studies, the following picture emerges. As shown previously, DNA strands containing only or mostly A and T bases can form stable *ps* duplexes with rWC pairing. In the case of the 16-mer DNA sequence we used, the stability of the *ps*-duplex is reduced by ca. 12°C relative to the *aps* duplex, or ca. 0.8°C per base pair. However, for RNA, 2'-OMe-RNA, p-RNA and homo-DNA formation of a *ps* duplex was not observed. In light of the characteristic backbone-base pair inclinations associated with each of the three oligonucleotide types, we may conclude that a substantial inclination will prevent self-pairing of single strands and formation of a *ps* duplex with rWC base pairs, irrespective of the tendencies of such strands to coil. However, we do not exclude the possibility of formation of a *ps* rWC paired duplex between a DNA and an

RNA strand. Indeed, evidence for the existence of such a duplex has been provided for a DNA dodecamer with incorporated isoguanosine and 5-methylisocytidine residues that pair with C and G, respectively, from the complementary RNA via three hydrogen bonds (56).

DNA and RNA can also self-pair and cross-pair by adopting a *ps* duplex with Hoogsteen base pairs, as inclination is not as restrictive in this case (see Table 1, triplex entry). Such duplexes are constituents of stable triple-stranded structures of variable DNA and RNA compositions (57, 58). However, not all combinations of DNA and RNA strands may exist: a DNA purine strand can pair with an RNA or a DNA pyrimidine complementary strand, but an RNA purine strand will only tolerate an RNA complement (46). The *ps* Hoogsteen-paired DNA duplex (59) is not to be confused with the aforementioned *ps* DNA duplex with rWC pairing (8) or the *aps* DNA duplex with Hoogsteen pairing (46). The structures of such *ps* Hoogsteen-paired duplexes between the all-purine strand 5'-GAAGGAAGA GAGAAAGGAGG and the all-pyrimidine strand 5'-CTT CCTTCTCTTTTCCTCC have been modeled for both

DNA and RNA (60). As in the *aps* WC B-DNA duplex and in the *ps* rWC DNA duplex, the average base pair-backbone inclinations in the Hoogsteen-paired *ps* DNA duplex seem to be small, but appear high for the *ps* Hoogsteen-paired RNA duplex (coordinates not available). Unlike in DNA (59) where experimental evidence has been provided for the isolated existence of such a *ps* duplex, more detailed data indicating formation of an isolated *ps* Hoogsteen-paired RNA duplex appear to be lacking to our knowledge. On the other hand, formation of a *ps* Hoogsteen-paired duplex between a homo-pyrimidine RNA strand and a homo-purine all-*Rp*-phosphorothioate DNA strand has recently been demonstrated (61).

Characterizing duplex conformations with η_B

In many protein–DNA complexes, the geometry of the DNA duplex differs from those of the canonical A- or B-forms (62,63). For example, CAP (catabolite gene activator protein) induces a 90° kink in the DNA duplex by binding in the major groove (64), and TBP (TATA box binding protein) causes unwinding and bending upon binding in the minor groove of the TATA box element (65,66). A distinct DNA conformation with an enlarged major groove and sharing both, features of A-DNA and B-DNA has been observed in a series of transcription factor–operator complexes, including several complexes of Zn-finger containing proteins (32,67). DNA wound around the histone proteins in the nucleosome core particle exhibits various geometries associated with bending, kinking, groove widening and narrowing as well as metal ion binding (68). We have calculated the backbone-base pair inclinations in the DNA complexes of CAP [(64); NDB code PDR006], TBP [(65); NDB code PDT012], Zif268, a Zn-finger protein [(67); NDB code PDT006] and the nucleosome core particle [(68); NDB code PD0287]. The results of the analysis are depicted in Figure 7. In the case of CAP and TBP, many of the local inclinations lie outside the standard A-DNA and B-DNA ranges. Strikingly, the inclinations in the DNA bound by the Zif268 protein are all intermediate between those of A-DNA and B-DNA, indicating the unique geometry of that duplex. While this does not provide a sufficient quantitative description of a duplex yet, it shows that η_B is sensitive to subtle conformational changes in duplexes as a consequence of their interactions with proteins and constitutes a useful new parameter for describing local and global helix geometries.

Inclination and thermodynamic stability

The dominance of inter-strand stacking over intra-strand stacking is one of the hallmarks of nucleic acid pairing systems with strongly inclined backbones such as homo-DNA and p-RNA (Figure 4). The earlier observation for p-RNA that 2'-overhanging bases (dangling ends) enhance the stability of duplexes significantly but 4'-dangling ends do not (36) can be easily rationalized with the negatively inclined backbone in this system.

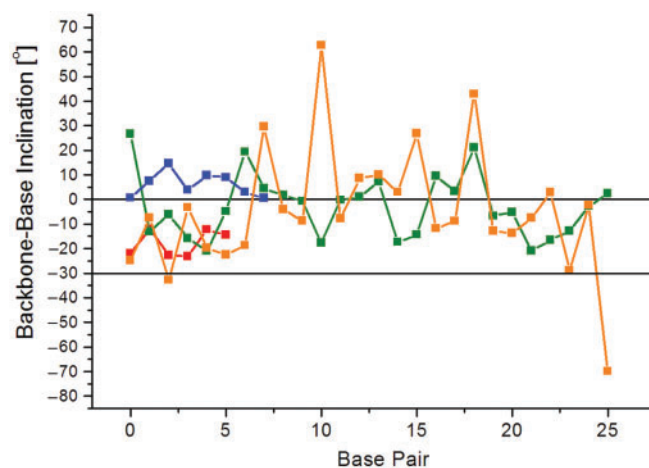


Figure 7. The inclination angle η_B is sensitive to distortions of the helix geometry. Base pair-backbone inclinations of duplex DNA in three transcription factor–DNA operator complexes, Zif268 (red, base pairs G4-C20 to C9-G15), CAP (green, base pairs A4-T60 to T29-A35), TBP (blue, base pairs A3-T27 to T10-A20), and the nucleosome core particle (orange, base pairs T71-A71 to A46-T46). The horizontal lines at 0° and –30° indicate the average inclination angles observed for B- and A-form duplexes, respectively. For a diagram depicting the inclination for all base pairs of the nucleosome core particle DNA please see the Supplementary Data.

On the other hand, only 6'-dangling ends would be expected to lead to an increase in stability in the homo-DNA series. To a smaller extent, the inclined backbone of RNA also leads to differential stabilization by 5'- and 3'-dangling ends, whereby the latter clearly display increased stability (69–71). Not surprisingly, the presence of 5'- versus 3'-dangling ends in DNA leads to no appreciable differences regarding stability (72) as the backbone is not inclined. The backbone inclination parameter is not only an important determinant of cross-pairing and relative strand polarity, but it also provides a measure for differential increases in the thermodynamic stability of nucleic acid duplexes.

CONCLUSIONS AND BIOLOGICAL SIGNIFICANCE

We have defined a parameter η_B that characterizes the inclination between the backbone and base planes in single-, double- and multi-stranded arrangements of nucleic acids. We have shown that this inclination is large in A-form DNA and RNA duplexes (ca. –30°) and that it is absent (ca. 0° on average) in B-form DNA. Whereas DNA can adopt parallel-stranded (*ps*) duplexes with rWC base-pairing, UV-melting studies have provided evidence that RNA does not. Similarly, the more linear homo-DNA and p-RNA pairing systems do not form a *ps* rWC duplex. The glucopyranose- and ribopyranose-based oligonucleotides display backbone inclinations of similar magnitude, but of opposite sign. The larger the backbone-base inclination in an oligonucleotide system, the lower its tendency to form a *ps* (rWC paired) duplex. This correlation appears to be independent of the various degrees of coiling inherent to different backbone types.

Therefore, we can exclude the *ps* rWC paired duplex as a possible folding motif in RNA molecules. Further, oligonucleotides that differ strongly in their relative backbone-base inclinations will not form *aps* duplexes with WC pairing either. For example, homo-DNA ($\eta_B = 35^\circ$) and p-RNA ($\eta_B = -46^\circ$) strands will not pair, but RNA ($\eta_B = -30^\circ$) and DNA ($\eta_B = 0^\circ$) pair because DNA adapts to the inclination of the former. Clearly, the relative degrees of twisting in the two strands will also affect their tendency to pair with each other. By comparison, standard helical parameters such as (helix axis-base pair) inclination or x - and y -displacement (shift and slide, respectively) (31) are not reliable indicators for the pairing behavior of nucleic acids. For example, in homo-DNA both the backbone-base (η_B) and the axis-base inclination (η) angles are strongly positive. In the case of RNA, η_B is negative and η is positive. Clearly, the positive inclinations between axis and base planes in both systems cannot serve as a reliable predictor of cross-pairing. Conversely, the opposite directions of backbone-base inclination in homo-DNA and RNA are consistent with the inability of the two systems to pair with one another.

Why are the chromosomes not composed of two DNA strands with parallel alignment? Although DNA can in principle adopt a *ps* rWC paired duplex, the *ps* structure has two main disadvantages compared with the standard *aps* one. The first disadvantage is its lower thermodynamic stability as a result of the formation of only two hydrogen bonds in the rWC G–C base pair compared with the three in the WC pairing mode. The second is the larger conformational rigidity of the *ps* duplex. The *aps* duplex features local pseudo-2-fold rotation axes, oriented normal to the helix axis and positioned within as well as halfway between the base-pair planes. In the *ps* duplex, the geometries of base pairs are all governed by one and the same global pseudo-2-fold rotation axis. With the rWC pairing mode, the base planes in the *ps* duplex are poised to be roughly perpendicular to the backbones and the helix axis. The adoption of a range of conformations as seen in the case of the *aps* WC paired arrangement is thus impossible (i.e. the DNA A-form with a negative backbone inclination).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR online and include UV melting curves for duplexes of 2'-OMe-modified RNA with *ps* and *aps* alignments, backbone-base inclination angles for the complete DNA in the nucleosome core particle crystal structure, and the README text for the Inclination program which provides a detailed description of the input and output file formats and the program output for a test structure. The source code can be compiled on any UNIX or LINUX machine and can be obtained from the authors upon request.

ACKNOWLEDGEMENTS

We thank the US National Institutes of Health for continued support (GM055237 to M.E.), Dr Stefan Pitsch for synthesis of the RNA and p-RNA oligonucleotides, Dr Christian Leumann for providing us with UV melting data for homo-DNA 16mers and Dr Bernhard Jaun for sending us the NMR coordinates of the p-RNA octamer duplex. Funding to pay the Open Access publication charges for this article was provided by NIH grant R01 GM055237.

Conflict of interest statement. None declared.

REFERENCES

1. Watson, J.D. and Crick, F.H.C. (1953) A structure for deoxyribose nucleic acid. *Nature*, **171**, 737–738.
2. Franklin, R.E. and Gosling, R.G. (1953) Molecular configuration in sodium thymonucleate. *Nature*, **171**, 740–741.
3. Pattabiraman, N. (1986) Can the double helix be parallel? *Biopolymers*, **25**, 1603–1606.
4. Van de Sande, J.H., Ramsing, N.B., Germann, M.W., Elhorst, W., Kalisch, B.W., von Kitzing, E., Pon, R.T., Clegg, R.C. and Jovin, T.M. (1988) Parallel stranded DNA. *Science*, **241**, 551–557.
5. Ramsing, N.B. and Jovin, T.M. (1988) Parallel stranded duplex DNA. *Nucleic Acids Res.*, **16**, 6659–6676.
6. Germann, M.W., Kalisch, B.W. and van de Sande, J.H. (1988) Relative stabilities of parallel- and antiparallel-stranded duplex DNA. *Biochemistry*, **27**, 8302–8306.
7. Rippe, K., Ramsing, N.B., Klement, R. and Jovin, T.M. (1990) A parallel stranded linear DNA duplex incorporating dG·dC base pairs. *J. Biomol. Struct. Dyn.*, **7**, 1199–1209.
8. Otto, C., Thomas, G.A., Jovin, T.M. and Peticolas, W.L. (1991) The hydrogen-bonding structure in parallel-stranded duplex DNA is reverse Watson-Crick. *Biochemistry*, **30**, 3062–3069.
9. Parvathy, V.R., Bhaumik, S.R., Chary, K.V.R., Govil, G., Liu, K., Howard, F.B. and Miles, H.T. (2002) NMR structure of a parallel-stranded DNA duplex at atomic resolution. *Nucleic Acids Res.*, **30**, 1500–1511.
10. Robinson, H., van der Marel, G.A., van Boom, J.H. and Wang, A.H.-J. (1992) Unusual DNA conformation at low pH revealed by NMR: parallel-stranded DNA duplex with homo base pairs. *Biochemistry*, **31**, 10510–10517.
11. Yang, X.-L., Sugiyama, H., Ikeda, S., Saito, I. and Wang, A.H.-J. (1998) Structural studies of a stable parallel-stranded DNA duplex incorporating isoguanine:cytosine and isocytosine:guanine basepairs by nuclear magnetic resonance spectroscopy. *Biophys. J.*, **75**, 1163–1171.
12. Gehring, K., Leroy, J.-L. and Guéron, M. (1993) A tetrameric DNA structure with protonated cytosine:cytosine base pairs. *Nature*, **363**, 561–565.
13. Chen, L., Cai, L., Zhang, X. and Rich, A. (1994) Crystal structure of a four-stranded intercalated DNA: d(C₄). *Biochemistry*, **33**, 13540–13546.
14. Kang, C., Berger, I., Lockshin, C., Ratliff, R., Moyzis, R. and Rich, A. (1995) Stable loop in the crystal structure of the intercalated four-stranded cytosine-rich metazoan telomere. *Proc. Natl. Acad. Sci. USA*, **92**, 3874–3878.
15. Berger, I., Kang, C., Fredian, A., Ratliff, R., Moyzis, R. and Rich, A. (1995) Extension of the four-stranded intercalated cytosine motif by adenine:adenine base pairing in the crystal structure of d(CCCAAT). *Nat. Struct. Biol.*, **2**, 416–425.
16. Eschenmoser, A. and Dobler, M. (1992) Why pentose and not hexose nucleic acids? Part I. Introduction to the problem, conformational analysis of oligonucleotide single strands containing 2',3'-dideoxyglucopyranosyl building blocks (homo-DNA), and reflections on the conformation of A- und B-DNA. *Helv. Chim. Acta*, **75**, 218–259.
17. Böhringer, M., Roth, H.-J., Hunziker, J., Göbel, M., Krishnan, R., Giger, A., Schweizer, B., Schreiber, J., Leumann, C. et al. (1992)

- Why pentose and not hexose nucleic acids? Part II. Preparation of oligonucleotides containing 2',3'-dideoxy- β -D-glucopyranosyl building blocks. *Helv. Chim. Acta*, **75**, 1416–1477.
18. Hunziker, J., Roth, H.-J., Böhringer, M., Giger, A., Diedrichsen, U., Göbel, M., Krishnan, R., Jaun, B., Leumann, C. *et al.* (1993) Why pentose and not hexose nucleic acids? Part III. Oligo(2',3'-dideoxy- β -D-glucopyranosyl) nucleotides ('homo-DNA'): base-pairing properties. *Helv. Chim. Acta*, **76**, 259–352.
 19. Groebke, K., Hunziker, J., Faser, W., Peng, L., Diedrichsen, U., Zimmermann, K., Holzner, A., Leumann, C. and Eschenmoser, A. (1998) Why pentose and not hexose nucleic acids? Part V. Purine-purine pairing in homo-DNA: guanine, isoguanine, 2,6-diaminopurine and xanthine. *Helv. Chim. Acta*, **81**, 375–474.
 20. Otting, G., Billeter, M., Wüthrich, K., Roth, H.-J., Leumann, C. and Eschenmoser, A. (1993) Why pentose and not hexose nucleic acids? Part IV. 'homo-DNA': ^1H -, ^{13}C -, ^{31}P - and ^{15}N -NMR-spectroscopic investigation of ddGlc(A-A-A-A-T-T-T-T) in aqueous solution. *Helv. Chim. Acta*, **76**, 2701–2756.
 21. Eschenmoser, A. (1991) Why pentose and not hexose nucleic acids? *Nachr. Chem. Tech. Lab.*, **39**, 795–806.
 22. Eschenmoser, A. and Loewenthal, E. (1992) Chemistry of potentially prebiological natural products. *Chem. Soc. Rev.*, **21**, 1–16.
 23. Eschenmoser, A. (1993) Toward a chemical etiology of the natural nucleic acids structure. In *Nouston, 40 Years of DNA Double Helix. Proceedings of the R. A. Welch Foundation* Nouston, TX, pp. 201–235.
 24. Pitsch, S., Wendeborn, S., Jaun, B. and Eschenmoser, A. (1993) Why pentose and not hexose nucleic acids? Part VII. Pyranosyl-RNA ('p-RNA'). *Helv. Chim. Acta*, **76**, 2161–2183.
 25. Pitsch, S., Krishnamurthy, R., Bolli, M., Wendeborn, S., Holzner, A., Minton, M., Lesueur, C., Schlönvogt, I., Jaun, B. *et al.* (1995) Pyranosyl-RNA ('p-RNA'): base pairing selectivity and potential to replicate. *Helv. Chim. Acta*, **78**, 1621–1635.
 26. Schlönvogt, I., Pitsch, S., Lesueur, C., Eschenmoser, A., Jaun, B. and Wolf, R.M. (1996) Pyranosyl-RNA (p-RNA): NMR and molecular-dynamics study of the duplex formed by self-pairing of ribopyranosyl-(C-G-A-A-T-T-C-G). *Helv. Chim. Acta*, **79**, 2316–2349.
 27. Eschenmoser, A. (1999) Chemical etiology of nucleic acid structure. *Science*, **284**, 2118–2124.
 28. Egli, M., Pallan, P.S., Pattanayek, R., Wilds, C.J., Lubini, P., Minasov, G., Dobler, M., Leumann, C.J. and Eschenmoser, A. (2006) Crystal structure of homo-DNA and nature's choice of pentose over hexose in the genetic system. *J. Am. Chem. Soc.*, **128**, 10847–10856.
 29. Yagil, G. and Sussman, J.L. (1986) Structural models of non-helical DNA. *EMBO J.*, **5**, 1719–1725.
 30. Saenger, W. (1984) *Principles of Nucleic Acid Structure* Springer Verlag Inc., New York, NY Springer Verlag Inc., New York, NY.
 31. Dickerson, R.E. (1989) Definitions and nomenclature of nucleic acid structure components. *Nucleic Acids Res.*, **17**, 1797–1803.
 32. Nekludova, L. and Pabo, C.O. (1994) Distinctive DNA conformation with enlarged major groove is found in Zn-finger-DNA and other protein-DNA complexes. *Proc. Natl Acad. Sci. USA*, **91**, 6948–6952.
 33. Arnott, S., Bond, P.J. and Chandrasekaran, R. (1980) Visualization of an unwound DNA duplex. *Nature*, **287**, 561–563.
 34. Olson, W.K. (1977) Spatial configuration of ordered polynucleotide chains: a novel double helix. *Proc. Natl Acad. Sci. USA*, **74**, 1775–1779.
 35. Watson, J.D. and Crick, F.H.C. (1953) The structure of DNA. *Cold Spring Harb. Symp. Quant. Biol.*, **18**, 123–131.
 36. Micura, R., Bolli, M., Windhab, N. and Eschenmoser, A. (1997) Pyranosyl-RNA also forms hairpin structures. *Angew. Chem. Int. Ed. Engl.*, **36**, 870–873.
 37. Lescrinier, E., Froeyen, M. and Herdewijn, P. (2003) Difference in conformational diversity between nucleic acids with a six-membered 'sugar' unit and natural 'furanose' nucleic acids. *Nucleic Acids Res.*, **31**, 2975–2989.
 38. Kennard, O. and Hunter, W.N. (1991) Single-crystal X-ray diffraction studies of oligonucleotides and oligonucleotide-drug complexes. *Angew. Chem. Int. Ed. Engl.*, **30**, 1254–1277.
 39. Newman, W.N. and Sproull, R.F. (1979) *Principles of Interactive Computer Graphics*, Chapter 21, 2nd edn. McGraw-Hill, New York.
 40. Berman, H.M., Olson, W.K., Beveridge, D.L., Westbrook, J., Gelbin, A., Demenz, T., Hsieh, S.H., Srinivasan, A.R. and Schneider, B. (1992) The Nucleic Acid Database: a comprehensive relational database of three-dimensional structures of nucleic acids. *Biophys. J.*, **63**, 751–759.
 41. Egli, M. (1994) Structural patterns in nucleic acids. In Bürgi, H.-B. and Dunitz, J.D. (eds), *Structure Correlation*, VCH Publishers Inc., New York/VCH Publishers Inc., New York, pp. 705–749.
 42. Haran, T.E., Shakked, Z., Wang, A.H.-J. and Rich, A. (1987) The crystal structure of d(CCCCGGGG): a new A-form variant with an extended backbone conformation. *J. Biomol. Struct. Dyn.*, **5**, 199–217.
 43. Portmann, S., Usman, N. and Egli, M. (1995) Crystal structure of r(CCCCGGGG) in two distinct lattices. *Biochemistry*, **34**, 7569–7575.
 44. Verdaguer, N., Aymami, J., Fernandez-Fornier, D., Fita, I., Coll, M., Huynh-Dinh, T., Igoien, J. and Subirana, J.A. (1991) Molecular structure of a complete turn of A-DNA. *J. Mol. Biol.*, **221**, 623–635.
 45. Bancroft, D., Williams, L.D., Rich, A. and Egli, M. (1994) The low-temperature crystal structure of the pure-spermine form of Z-DNA reveals binding of a spermine molecule in the minor groove. *Biochemistry*, **33**, 1073–1086.
 46. Abrescia, N., Thompson, A., Huynh-Dinh, T. and Subirana, J.A. (2002) Crystal structure of an antiparallel DNA fragment with Hoogsteen base pairing. *Proc. Natl Acad. Sci. USA*, **99**, 2806–2811.
 47. Nowotny, M., Gaidamakov, S.A., Crouch, R.J. and Yang, W. (2005) Crystal structures of RNase H bound to an RNA/DNA hybrid: substrate specificity and metal-dependent catalysis. *Cell*, **121**, 1005–1016.
 48. Laughlan, G., Murchie, A.I., Norman, D.G., Moore, M.H., Moody, P.C., Lilley, D.M. and Luisi, B.F. (1994) The high-resolution crystal structure of a parallel-stranded guanine tetraplex. *Science*, **265**, 520–524.
 49. Haider, S., Parkinson, G.N. and Neidle, S. (2002) Crystal structure of the potassium form of an *Oxytricha nova* G-quadruplex. *J. Mol. Biol.*, **320**, 189–200.
 50. Rasmussen, H., Kastrop, J.S., Nielsen, J.N., Nielsen, J.M. and Nielsen, P.E. (1997) Crystal structure of a peptide nucleic acid (PNA) duplex at 1.7 Å resolution. *Nat. Struct. Biol.*, **4**, 98–101.
 51. Menchise, V., De Simone, G., Tedeschi, T., Corradini, R., Sforza, S., Marchelli, R., Capasso, D., Saviano, M. and Pedone, C. (2003) Insights into peptide nucleic acid (PNA) structural features: the crystal structure of a d-lysine-based chiral PNA–DNA duplex. *Proc. Natl Acad. Sci. USA*, **100**, 12021–12026.
 52. Radhakrishnan, I. and Patel, D.J. (1993) Solution structure of a purine-purine•pyrimidine DNA triplex containing G-G•C and T-A•T triples. *Structure*, **1**, 135–152.
 53. Lesnik, E.A., Guinasso, C.J., Kawasaki, A.M., Sasmor, H., Zounes, M., Cummins, L.L., Ecker, D.J., Cook, P.D. and Freier, S.M. (1993) Oligodeoxynucleotides containing 2'-O-modified adenosine: synthesis and effects on stability of DNA:RNA duplexes. *Biochemistry*, **32**, 7832–7838.
 54. Lubini, P., Zürcher, W. and Egli, M. (1994) Crystal structure of a DNA duplex containing 2'-O-methylated adenosines. *Chem. Biol.*, **1**, 39–45.
 55. Manoharan, M. (2004) RNA interference and chemically modified small interfering RNAs. *Curr. Opin. Chem. Biol.*, **8**, 570–579.
 56. Seela, F., He, Y. and Wei, C. (1999) Parallel-stranded oligonucleotide duplexes containing 5-methylisocytosine-guanine and isoguanine-cytosine base pairs. *Tetrahedron*, **55**, 9481–9500.
 57. Roberts, R.W. and Crothers, D.M. (1992) Stability and properties of double and triple helices: dramatic effects of RNA and DNA backbone composition. *Science*, **258**, 1463–1466.
 58. Han, H. and Dervan, P.B. (1993) Sequence-specific recognition of double helical RNA and RNA:DNA by triple helix formation. *Proc. Natl Acad. Sci. USA*, **90**, 3806–3810.

59. Liu, K., Miles, T.H., Frazier, J. and Sasisekharan, V. (1993) A novel DNA duplex. A parallel-stranded DNA helix with Hoogsteen base pairing. *Biochemistry*, **32**, 11802–11809.
60. Raghunathan, G., Miles, H.T. and Sasisekharan, V. (1994) Parallel nucleic acid helices with Hoogsteen base pairing: symmetry and structure. *Biopolymers*, **34**, 1573–1581.
61. Guga, P., Janicka, M., Maciaszek, A., Rebowska, B. and Nowak, G. (2007) Hoogsteen paired homopurine (Rp-PS)-DNA and homopyrimidine RNA strands form a thermally stable parallel duplex. *Biophys. J.*, in press.
62. Steitz, T.A. (1993) *Structural Studies of Protein-Nucleic Acid Interaction*. Cambridge University Press, Cambridge, UK
Cambridge University Press, Cambridge, UK.
63. Luisi, B. (2006) Protein-nucleic acid interactions. In Blackburn, G.M., Gaits, M.J., Loakes, D. and Williams, D.M. (eds), *Nucleic Acids in Chemistry and Biology, 3rd edn.* RSC Publishing, Cambridge, UK; RSC Publishing, Cambridge, UK, pp. 383–425.
64. Schultz, S.C., Shields, G.C. and Steitz, T.A. (1991) Crystal structure of a CAP-DNA complex. The DNA is bent by 90°. *Science*, **253**, 1001–1007.
65. Kim, Y., Geiger, J.H., Hahn, S. and Sigler, P.B. (1993) Crystal structure of a yeast TBP/TATA-box complex. *Nature*, **365**, 512–520.
66. Kim, J.L., Nikolov, D.B. and Burley, S.K. (1993) Co-crystal structure of TBP recognizing the minor groove of a TATA element. *Nature*, **365**, 520–527.
67. Pavletich, N.P. and Pabo, C.O. (1991) Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å. *Science*, **252**, 809–817.
68. Davey, C.A. and Richmond, T.J. (2003) The structure of DNA in the nucleosome core. *Nature*, **423**, 145–150.
69. Petersheim, M. and Turner, D.H. (1983) Base-stacking and base-pairing contributions to helix stability: thermodynamics of double-helix formation with CCGG, CCGGp, CCGGAp, ACCGGp, ACCGGUp, and ACCGGUp. *Biochemistry*, **22**, 256–263.
70. Freier, S.M., Burger, B.J., Alkema, D., Neilson, T. and Turner, D.H. (1983) Effects of 3' dangling end stacking on the stability of GGCC and CCGG double helices. *Biochemistry*, **22**, 6198–6206.
71. Freier, S.M., Kierzek, R., Jaeger, J.A., Sugimoto, N., Caruthers, M.H., Neilson, T. and Turner, D.H. (1986) Improved free-energy parameters for predictions of RNA duplex stability. *Proc. Natl Acad. Sci. USA*, **83**, 9373–9377.
72. Senior, M., Jones, R.A. and Breslauer, K.J. (1988) Influence of dangling thymidine residues on the stability and structure of two DNA duplexes. *Biochemistry*, **27**, 3879–3885.