J|A|C|S
ARTICLES
Published on Web 08/01/2006

Crystal Structure of Homo-DNA and Nature's Choice of Pentose over Hexose in the Genetic System

Martin Egli,*,† Pradeep S. Pallan,† Rekha Pattanayek,† Christopher J. Wilds,†,‡ Paolo Lubini,§ George Minasov,|| Max Dobler,⊥ Christian J. Leumann,# and Albert Eschenmoser+,@

Contribution from the Department of Biochemistry, School of Medicine, Vanderbilt University, Nashville, Tennessee 37232, Department of Chemistry and Biochemistry, Concordia University, Montreal, Quebec H4B 1R6, Canada, Alta Scuola Pedagogica, CH-6600 Locarno, Switzerland, Department of Molecular Pharmacology and Biological Chemistry, School of Medicine, Northwestern University, Chicago, Illinois 60611, Biographics Laboratory 3R, CH-4056 Basel, Switzerland, Department of Chemistry and Biochemistry, University of Bern, CH-3012 Bern, Switzerland, Department of Chemistry and The Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, California 92037, and Laboratory of Organic Chemistry, ETH Hönggerberg, CH-8093 Zürich, Switzerland

Received April 20, 2006; E-mail: martin.egli@vanderbilt.edu

Abstract: An experimental rationalization of the structure type encountered in DNA and RNA by systematically investigating the chemical and physical properties of alternative nucleic acids has identified systems with a variety of sugar−phosphate backbones that are capable of Watson−Crick base pairing and in some cases cross-pairing with the natural nucleic acids. The earliest among the model systems tested to date, (4′ → 6′)-linked oligo(2′,3′-dideoxy-β-D-glucopyranosyl)nucleotides or homo-DNA, shows stable self-pairing, but the pairing rules for the four natural bases are not the same as those in DNA. However, a complete interpretation and understanding of the properties of the hexapyranosyl (4′ → 6′) family of nucleic acids has been impeded until now by the lack of detailed 3D-structural data. We have determined the crystal structure of a homo-DNA octamer. It reveals a weakly twisted right-handed duplex with a strong inclination between the hexose−phosphate backbones and base-pair axes, and highly irregular values for helical rise and twist at individual base steps. The structure allows a rationalization of the inability of allo-, altro-, and glucopyranosyl-based oligonucleotides to form stable pairing systems.

Introduction

Chemical synthesis of alternative nucleic acid-pairing systems and an exploration of their physical and chemical properties can potentially yield insights into nature's choice of pentoses over other potential candidates as the carbohydrate building blocks of the genetic material (conceptualized and reviewed in refs 1 and 2; for selected examples, see refs 3−7). The question "why pentose and not hexose?" marked the beginning of a systematic investigation of an etiology of nucleic acid structure.[8,9] Detailed studies of the pairing properties of oligo(2′,3′-dideoxy-β-D-glucopyranosyl)nucleotides (homo-DNA; Figure 1) demonstrated that DNA is not unique in regard to the formation of duplexes with Watson−Crick base pairing.[10,11] Theoretical considerations of the conformational preferences of homo-DNA pointed to only two possible combinations of the six backbone torsion angles that would generate a repetitive arrangement. This led to the view that the helicality of DNA is a consequence of the inherent geometrical constraints of the deoxyriboses in its backbone.[12] Furthermore, the Watson−Crick base-pairing priorities in DNA duplexes (G−C > A−T) are no longer valid in duplexes formed by homo-DNA (G−C > A−A ≈ G−G >

† Vanderbilt University.
‡ Concordia University.
§ Alta Scuola Pedagogica.
|| Northwestern University.
⊥ Biographics Laboratory 3R.
University of Bern.
+ The Scripps Research Institute.
@ ETH Hönggerberg.

(1) Eschenmoser, A.; Kisakürek, M. V. *Helv. Chim. Acta* **1996**, *79*, 1249−1259.
(2) Eschenmoser, A. *Science* **1999**, *284*, 2118−2124.
(3) Pitsch, S.; Krishnamurthy, R.; Bolli, M.; Wendeborn, S.; Holzner, A.; Minton, M.; Lesueur, C.; Schlönvogt, I.; Jaun, B.; Eschenmoser, A. *Helv. Chim. Acta* **1995**, *78*, 1621−1635.
(4) Beier, M.; Reck, F.; Wagner, T.; Krishnamurthy, R.; Eschenmoser, A. *Science* **1999**, *283*, 699−703.
(5) Jungmann, O.; Wippo, H.; Stanek, M.; Huynh, H. K.; Krishnamurthy, R.; Eschenmoser, A. *Org. Lett.* **1999**, *1*, 1527−1530.
(6) Reck, F.; Wippo, H.; Kudick, R.; Bolli, M.; Ceulemans, G.; Krishnamurthy, R.; Eschenmoser, A. *Org. Lett.* **1999**, *1*, 1531−1534.
(7) Schöning, K. U.; Scholz, P.; Guntha, S.; Wu, X.; Krishnamurthy, R.; Eschenmoser, A. *Science* **2000**, *290*, 1347−1351.
(8) Eschenmoser, A.; Loewenthal, E. *Chem. Soc. Rev.* **1992**, *21*, 1−16.
(9) Eschenmoser, A. In *40 years of the DNA double helix*; Proceedings of the Robert A. Welch Foundation Conference on Chemical Research; Houston, TX, Oct 25 and 26, 1993; R. A. Welch Foundation: Houston, TX, 1993; pp 201−235.
(10) Böhringer, M.; Roth, H.-J.; Hunziker, J.; Göbel, M.; Krishnan, R.; Giger, A.; Schweizer, B.; Schreiber, J.; Leumann, C.; Eschenmoser, A. *Helv. Chim. Acta* **1992**, *75*, 1416−1477.
(11) Hunziker, J.; Roth, H.-J.; Böhringer, M.; Giger, A.; Diedrichsen, U.; Göbel, M.; Krishnan, R.; Jaun, B.; Leumann, C.; Eschenmoser, A. *Helv. Chim. Acta* **1993**, *76*, 259−352.
(12) Eschenmoser, A.; Dobler, M. *Helv. Chim. Acta* **1992**, *75*, 218−259.

10.1021/ja062548x CCC: $33.50

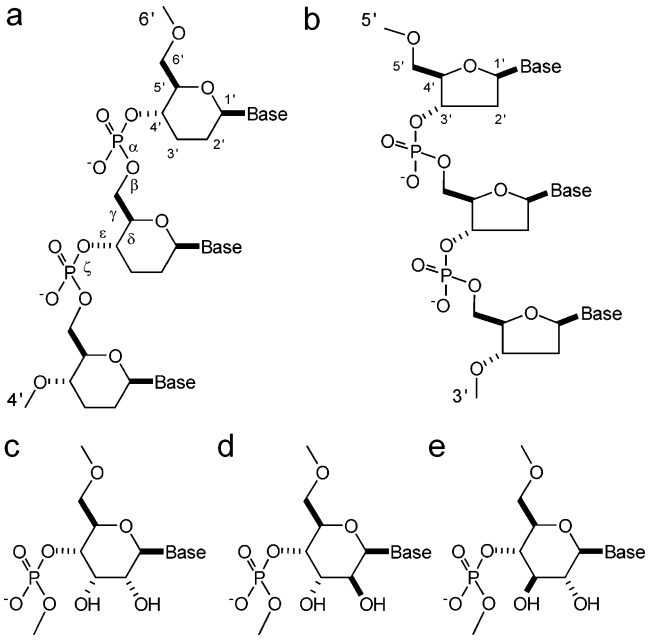

Figure 1. Structures, configurations, and linkage mode of natural and artificial oligonucleotide systems. (a) Homo-DNA (2′,3′-dideoxy-β-D-glucopyranose sugars), (B) DNA, (C) β-D-allopyranosylnucleotide, (d) β-D-altropyranosylnucleotide, and (e) β-D-glucopyranosylnucleotide. Sugar atoms in homo-DNA and DNA are numbered; the conformation of the backbone for both can be described by six torsion angles α to ζ.

A−T).[11] Thus, the relative stability of the standard base pairs in DNA is related to the nature of the backbone sugar moiety, and the 2′,3′-dideoxyglucopyranose appears to facilitate purine−purine pairing via the antiparallel reverse-Hoogsteen mode.[11,13]

Another characteristic of homo-DNA is the more favorable contribution of entropy to pairing compared with DNA and the strictly antiparallel orientation of strands.[11] The former can be explained with the higher rigidity of the hexose sugar relative to pentose, and the latter is consistent with an idealized linear model of the duplex in which the backbones are inclined relative to base-pair planes.[11] Homo-DNA constitutes an autonomous pairing system and does not pair with DNA or any of the artificial nucleic acid systems investigated thus far. It is noteworthy that this behavior differs from that of α-homo-DNA that pairs with RNA under formation of a parallel-oriented non-A-, non-B-type duplex structure,[14] and another pairing system based on a hexose sugar, hexitol nucleic acid (HNA), that adopts an A-form duplex conformation[15] and pairs with RNA. The term homo-DNA used in this article shall refer exclusively to β-homo-DNA.

The conformational properties of homo-DNA single and double strands have been the focus of numerous studies. Models based on backbone torsion angles with ideal *synclinal* (*sc*) or *antiperiplanar* conformations (*ap*) (the hexose assumes a chair conformation with all three substituents in the equatorial orientation; Figure 1a) are essentially linear.[12] Conformational variants with α/γ pairs in either the sc^-/sc^+ or *ap*/*ap* conformations ($\chi = -120°$) lead to a distance between adjacent bases of ca. 5 and 6 Å, respectively. Molecular modeling of a homo-DNA single strand showed a weakly twisted right-handed conformation with a helical pitch of ca. 120−130 Å (see footnote 20 in ref 12). NMR solution experiments in combination with modeling for a self-complementary homo-DNA duplex of sequence A_5T_5 resulted in two more or less linear models with angles of ca. 60 and 45° between the strand directions and base-pair axes (corresponding to the two above torsion angle variants) and distances between adjacent base pairs along their normals of 4.5 Å and more.[16] A recent molecular dynamics simulation suggested a helical right-handed conformation for the homo-DNA duplex with a maximum twist of 10° (36 base pairs per turn).[17] Nevertheless, such models exhibit a distance between adjacent base pairs that is considerably larger than the ideal value of 3.4 Å and are thus incompatible with effective stacking. Therefore, the lack of a high-resolution structure for homo-DNA leaves many questions unanswered. For example, a definitive answer as to why the β-D-allo-, β-D-altro-, and β-D-glucopyranosyl nucleic acid systems do not display pairing (ref 2 and references cited therein; Figure 1c−e) cannot be given without more detailed structural data.

Crystals of a homo-DNA duplex were available as early as 1992, but all attempts to phase the initial and subsequently improved diffraction data had failed over the years. We have now determined the structure using a strategy specifically developed for solving the phase problem with difficult oligonucleotide crystal structures.[18] Here we describe the structure of the homo-DNA duplex $[dd(CGAATTCG)]_2$ (dd = 2′,3′-*did*eoxy-β-D-glucopyranose sugars) and insights regarding the role of backbone−base inclination and interstrand stacking in pairing selectivity. Our results also allow a rationalization for the absence of pairing with (4′ → 6′)-linked fully hydroxylated hexopyranosyl nucleic acids and a refined answer to the question "why pentose and not hexose sugars?", raised as part of an experimental investigation of an etiology of nucleic acid structure.

Results

Structure Determination Attempts. To determine the crystal structure of a homo-DNA duplex, sequences of varying lengths (2−12 nucleotides long), some of them involving purine−purine pairs, were subjected to crystallization trials. Despite numerous attempts and using more standard approaches for crystallization of oligonucleotides (i.e., polyamines and 2-methyl-2,4-pentanediol, MPD) as well as a variety of commercially available sparse matrix screens, crystals could only be grown for the octamer dd(CGAATTCG). These crystals were obtained from solutions containing magnesium chloride, sodium cacodylate buffer pH 7, and MPD as the precipitant. Variations of the conditions, such as, for example, replacement of Mg^{2+} by Ca^{2+} or Na^+ by K^+, or addition of spermine, did not produce crystals. The crystals have the appearance of thick hexagonal rods and belong to the enantiomorphic space group pair $P6_122/P6_522$. Initial diffraction data collected on in-house rotating anode sources exhibited maximum resolution limits of around 2.7 Å.

(13) Groebke, K.; Hunziker, J.; Faser, W.; Peng, L.; Diedrichsen, U.; Zimmermann, K.; Holzner, A.; Leumann, C.; Eschenmoser, A. *Helv. Chim. Acta* **1998**, *81*, 375−474.
(14) Froeyen, M.; Lescrinier, E.; Kerremans, L.; Rosemeyer, H.; Seela, F.; Verbeure, B.; Lagoja, I.; Rozenski, J.; Van Aerschot, A.; Busson, R.; Herdewijn, P. *Chem.−Eur. J.* **2001**, *7*, 5183−5194.
(15) Declercq, R.; Van Aerschot, A.; Read, R. J.; Herdewijn, P.; Van Meervelt, L. *J. Am. Chem. Soc.* **2002**, *124*, 928−933.
(16) Otting, G.; Billeter, M.; Wüthrich, K.; Roth, H.-J.; Leumann, C.; Eschenmoser, A. *Helv. Chim. Acta* **1993**, *76*, 2701−2756.
(17) Lescrinier, E.; Froeyen, M.; Herdewijn, P. *Nucleic Acids Res.* **2003**, *31*, 2975−2989.
(18) Wilds, C. J.; Pattanayek, R.; Pan, C.; Wawrzak, Z.; Egli, M. *J. Am. Chem. Soc.* **2002**, *124*, 14910−14916.

Table 1. Phosphoroselenoate MAD Phasing Statistics and Selected Native Data Collection and Refinement Parameters

MAD data collection and phasing statistics				native data statistics and refinement parameters	
space group	hexagonal $P6_122$			space group	hexagonal $P6_122$
unit cell	$a = b = 38.9\ 6$ Å, $c = 134.16$ Å			unit cell	$a = b = 38.94$ Å, $c = 133.85$ Å
wavelength	0.9797 Å inflection	0.9794 Å peak	0.9701 Å remote	wavelength	1.000 Å
temperature	−160 °C			temperature	−160 °C
resolution	50−2.10 Å (last shell 2.18−2.10 Å)			resolution	1.75 Å
redundancy	12.4 (12.8)	12.4 (12.5)	6.2 (6.5)	total reflections collected	86414
unique data	3981 (382)	3997 (385)	3967 (382)	unique data	6688
completeness	99.4% (100)	99.5% (100)	99.3% (100)	completeness (1.81−1.75 Å)	99.2% (98.9)
R_{merge}[a]	0.076 (0.457)	0.089 (0.533)	0.078 (0.516)	R_{merge}[a] (1.81−1.75 Å)	0.035 (0.22)
R_{Cullis}[b]				R[d] (all reflections)	0.239
centric/ acentric	1.35/1.32	0.80/0.90		R_{work}[d]	0.240
all	1.32	0.87		R_{free}[e]	0.282
phasing power[c]				no. of water molecules	70
centric/ acentric	2.46/2.13	2.93/2.82		no. of metal ions	1 Mg^{2+}
all	2.20	2.84		rmsd bond lengths	0.015 Å
figure of merit				rmsd bond angles	1.2°
centric/ acentric	0.54/0.25	0.53/0.28		ave B-factor DNA atoms	41 Å^2
all	0.30	0.32		ave B-factor water molecules	50 Å^2

[a] $R_{merge} = \Sigma_{hkl}\Sigma_i|I(hkl)_i - \langle I(hkl)\rangle|/\Sigma_{hkl}\Sigma_i\langle I(hkl)_i\rangle$ for i measurements of the intensity I of a reflection hkl. [b] $R_{Cullis} = \Sigma||F_{(\lambda i)} \pm F_{(\lambda 1)}| - |F_{h(\lambda i),c}||/\Sigma|F_{(\lambda i)} \pm F_{(\lambda 1)}|$, where $F_{h\ (\lambda i),c}$ is the calculated heavy structure factor. [c] Phasing power $= \langle F_{h(\lambda i)}\rangle/E$, where $\langle F_{h(\lambda i)}\rangle$ is the rms heavy atom structure factor and E is the residual lack-of-closure error. [d] $R = \Sigma_{hkl}||F_{ohkl}| - k|F_{chkl}||/\Sigma_{hkl}\ |F_{ohkl}|$, where $|F_{ohkl}|$ and $|F_{chkl}|$ are the observed and calculated structure factor amplitudes, respectively. [e] R_{free} idem, for a set of reflections (5% of the total) omitted from the refinement process.

However, data subsequently collected on an insertion device beamline at the Advanced Photon Source are almost complete to 1.75 Å resolution (Table 1).

The crystals and diffraction data were analyzed exhaustively (including experimental density measurements and calculations of self-rotation function and Patterson maps), and numerous attempts were undertaken to determine the structure by the molecular replacement technique. However, none of the trial models used in the rotation and translation searches, including those based on ideal conformations of backbone torsion angles[12] and other duplex structures of varying helical rise and twist that were also subjected to molecular dynamics simulations and energy minimization, appeared to be even close to the conformation of the octamer in the crystal. We synthesized the Br^5U and Br^5C homo-DNA phosphoramidite building blocks, but none of the four derivatized strands (Br^5CGAATTCG, CGAABr^5UTCG, CGAATBr^5UCG, and CGAATTBr^5CG) produced crystals under either the conditions used to grow crystals of the native octamer or variations thereof. Cocrystallization experiments with or soaking in solutions of a battery of heavy atom compounds including alkali (Rb^+, Cs^+) and alkaline earth (Sr^{2+}, Ba^{2+}) metal ions as well as exposure of octamer crystals to xenon in a pressure cell all failed to produce derivative crystals. It was concluded that the hexagonal crystals most likely contained two octamer strands per asymmetric unit. However, it was not clear whether the strands belonged to two independent duplexes, both located on twofold rotation axes, or a single duplex in a general position. Thus, after numerous attempts to crack the phase problem with these crystals and lacking another crystal form or crystals of other sequences, structure determination appeared to be at a dead end.

Phasing with a Single Phosphoroselenoate Derivative. The project lay dormant for several years until we considered the use of phosphorothioates and phosphoroselenoates (PS and PSe, respectively) for determination of oligonucleotide crystal structures. The former are chemically stable, whereas PSe-DNAs were considered too unstable for X-ray crystallographic applications due to oxidation to the phosphate form in a matter of hours or a few days. To test the suitability of PS-DNA for phasing purposes, we produced all 10 diastereoisomerically pure CGCGCG hexamers with a single PS moiety per strand (Wilds and Egli, unpublished data). Crystals could be grown for most of them, but diffraction data collected at low-energy wavelengths on synchrotron beamlines exhibited only weak anomalous effects based on the two sulfur atoms per duplex, unsuitable for phasing by the single or multiple wavelength anomalous dispersion techniques (SAD or MAD, respectively). Attempts to derivatize PS-DNA crystals with Tl(I) or Hg(II) compounds all failed, leading to disintegration of crystals or poor diffraction data with high mosaicity.

Subsequently, we found that single PSe moieties per hexamer allowed phasing of d(CGCGCG) crystals.[18] However, the yields of chemical syntheses of PSe-DNAs were rather poor. Placement of the PSe moiety near the 3′-end poses a challenge due to the considerable loss of the functionality during the oxidative step of each single-nucleotide extension. Further reductions in yield are incurred by the need to separate diastereoisomeric pairs of oligonucleotides using ion exchange chromatography (see Materials and Methods). Thus, synthesis of all 14 possible diastereoisomerically pure dd(CGAATTCG) strands with single PSe moieties was out of the question as syntheses of the required amounts of phosphoramidite building blocks would have been prohibitively expensive. However, we produced all 14 PS homo-DNA octamers in the expectation that crystallization trials with these would allow a good prediction as to which of the PSe homo-DNAs would still yield to crystallization (data not shown; selenium has a van der Waals radius of 2 Å compared to 1.85 Å for sulfur and 1.4 Å for oxygen). Surprisingly, only two of the 14 PS homo-DNAs produced crystals. These correspond to the more slowly eluting diastereoisomers of the dd(CGA_{PS}-ATTCG) and dd($CGAATTC_{PS}G$) pairs in the strong anion exchange HPLC purifications (peak 2, see Figure S1, Supporting Information). The absolute configuration of the PS moiety, R_P versus S_P, could not be assigned at that point. On the basis of these observations, we decided to synthesize sufficient amounts of the peak-2 octamer dd(CGA_{PSe}ATTCG) for crystallization

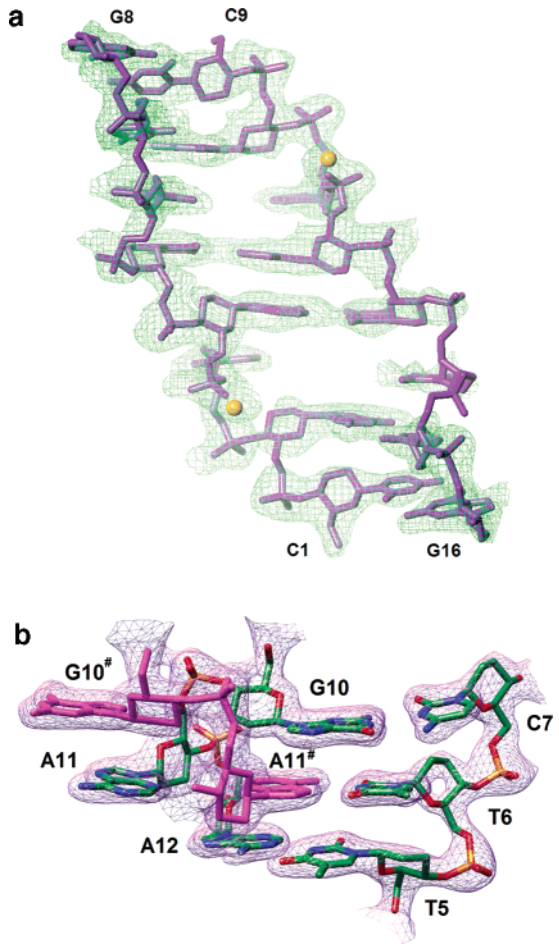

Figure 2. Phosphoroselenoate MAD phasing and quality of the final model. (a) Experimental map following density modification superimposed on the final structure. Selenium atoms are shown as yellow spheres. (b) ($2F_o - F_c$) Sum electron density based on the final model and drawn at the 1σ level. All residues are numbered, and a GpA dinucleotide from a symmetry-related molecule is shown with bonds in magenta.

and structure determination. The other octamer was considered unsuitable due to the location of the PSe moiety close to the 3′-terminus and the resulting challenges for chemical synthesis.

Three different attempts were made to collect optimal MAD data for a dd(CGA$_{PSe}$ATTCG) crystal. The final success was critically dependent on careful timing of synthesis and synchrotron data collection. The diffraction data for which a summary is shown in Table 1 were acquired a week after the HPLC separation of the two PSe homo-DNA diastereoisomers. For MAD phasing with the program CNS[19] the resolution of the data was limited to 2.1 Å as the normalized χ^2 value calculated in HKL-2000[20] fell below 1.5 for reflection shells with resolutions <2.1 Å. The resulting density-modified experimental map in space group $P6_122$ was of excellent quality (Figure 2a), and all 16 nucleotides could be readily placed into the electron density. The configuration of the phosphoroselenoate moiety is S_P. The initially built DNA model resulted in an *R*-factor of 44%, and refinement with the program CNS lowered the R_{work} to 34%. Further refinement with an all-chair homo-DNA model in REFMAC[21] using native data to 1.75 Å resolution and treating homo-DNA atoms with anisotropic temperature factors resulted in a final *R*-factor of 23%. An example of the quality of the ($2F_o - F_c$) sum electron density is depicted in Figure 2b, and selected refinement parameters are listed in Table 1. The solvent content of homo-DNA crystals is 55%. However, the current model comprises only 70 water molecules, and the relatively high values for R_{work} and R_{free} are to some extent a result of the poorly defined electron density in the large solvent channels.

Homo-DNA Duplex Geometry and Sequence Dependence of Conformation. In the homo-DNA crystal, octamers are paired in an antiparallel fashion under formation of Watson−Crick pairs. One base per strand is extruded from the duplex: A3 in strand 1 and A11 in strand 2 (Figure 3a; nucleotides are numbered 1−8 and 9−16 in strands 1 and 2, respectively). Adenines from a symmetry-related duplex pair with T's in the reverse-Hoogsteen mode at these sites (Figure 4). The duplex has overall dimensions of 40 × 24 × 24 Å and is more compact than the anticipated linear models. It has a right-handed twist that amounts to 14° on average per base-pair step (Table 2). However, individual steps exhibit highly irregular twists, and in some of them no twisting occurs at all (Figure 3b). The accumulated twist between G2 and A4 amounts to about 45°, and another high twist is observed between T5 and T6 (32°).

The average rise is 3.8 Å, but the values also vary considerably at individual base steps (Table 2). In addition, the distance along the normal to base planes between intrastrand bases differs somewhat from the stacking distance between bases from opposite strands. An example of this is seen at the central ApT step of the duplex. There, the rise between A4 and T5 (similar for A12 and T13) is 4.1 Å, but A4 and A12 are stacked at an ideal distance of 3.4 Å under formation of a minimal twist (Figure 3b). This pure cross-strand stacking is brought about by a base slide of almost 6 Å (*y*-displacement; Table 2), and no overlap occurs between A's and T's. A similarly large slide without twisting is seen with base-pair steps at both duplex ends. However, the rise is somewhat larger at these sites and is accompanied by significant rolling. The latter feature is most likely a consequence of packing interactions that involve stacking between terminal base pairs from symmetry-related duplexes.

Although the geometries of individual base-pair steps in homo-DNA vary considerably, a feature shared by all of them is the virtual absence of intrastrand stacking. This is particularly evident at pyrimidine−purine (C1pG2, C7pG8) and purine−pyrimidine (A4pT5) steps that all exhibit large slides, practically no twist, and extensive overlaps between bases from opposite strands. Another common property is the surprisingly short distance between adjacent intrastrand phosphates (ave 5.8 Å; Table 2) that is comparable to that seen in A-form RNA duplexes.[22] The structure demonstrates that hexose sugars in homo-DNA do not result in a larger separation of adjacent nucleotides than in DNA duplexes. In fact, the average phosphate−phosphate spacing is clearly below that seen in B-form DNA (ca. 7 Å).

Crystal Packing and Interduplex Base Swapping. The homo-DNA duplex lacks a major groove and features a convex

(19) Brünger A. T.; Adams P. D.; Clore G. M.; DeLano W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J. S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L. *Acta Crystallogr., Sect. D* **1998**, *54*, 905−921.
(20) Otwinowski, Z.; Minor, W. *Methods Enzymol.* **1997**, *276*, 307−326.
(21) Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. *Acta Crystallogr., Sect. D* **1997**, *53*, 240−255.
(22) Rich, A. *Nat. Struct. Biol.* **2003**, *10*, 247−249.

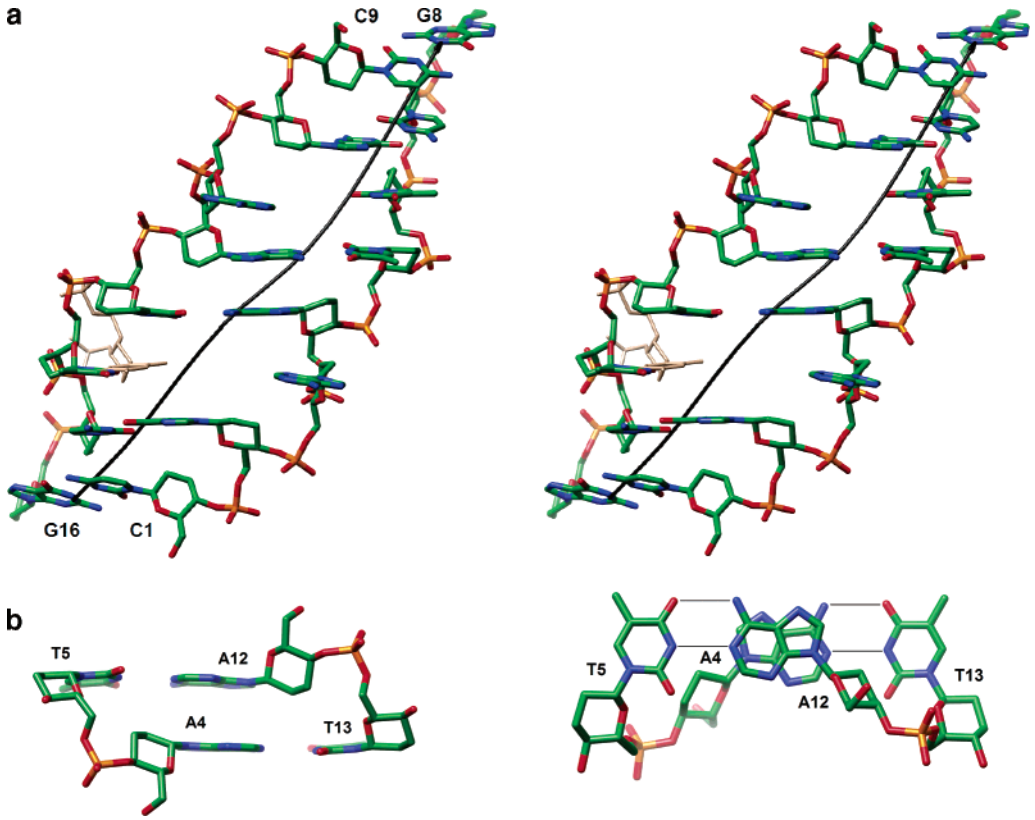

Figure 3. Geometry of the homo-DNA octamer duplex. (a) Stereo diagram of a view onto the convex surface and approximately along the molecular dyad. Atoms are colored green, red, blue, and orange for carbon, oxygen, nitrogen, and phosphorus, respectively, and residues of terminal base pairs are labeled. The helix axis is shown as a black line and was calculated by replacing the positions of looped-out A3 and A11 with bases from a symmetry-related duplex that form reverse-Hoogsteen pairs with T6 and T14. An alternative backbone conformer with a chair conformation of the hexose in the region of T14 is shown with thin bonds in beige. (b) Close-up views of the central ApT base-pair step along the molecular dyad into the minor groove (left), and rotated by 90° and normal to base pairs (right). The views illustrate the absence of a twist and an ideal 3.4 Å stacking distance at this step.

surface instead. The minor groove is shallow and about 11 Å wide on average. Thus, the duplex resembles a slowly writhing ribbon, distinct from the familiar cylindrical shapes of the right- and left-handed DNA duplexes (Figure 4a). Although the information contents are the same, the convex surfaces of homo-DNA and Z-DNA have nothing in common topologically.[23] Not only do they exhibit different handedness but also the surface presented by homo-DNA is much less wound (Figure 4).

In the crystal, homo-DNA duplexes dimerize around a crystallographic dyad via hydrogen bond acceptors and donors exposed on their convex surfaces (Figures 4 and 5). Duplexes cross at an angle of ca. 60° and are intertwined so tightly that Watson−Crick base pairs are disrupted at two locations to avoid a clash (A3-T14 and T6-A11; Figures 3 and 4). The gap left by looped-out adenines is filled by adenines form the paired duplex, whereby A's interact with T's in a reverse-Hoogsteen mode (Figure 5c). To our knowledge, this is the first occurrence of base swapping in the structure of a nucleic acid duplex. A recently reported crystal structure of an RNA−DNA hybrid exhibited swapping of base pairs.[24] The crossover of homo-DNA duplexes locally generates a four-stranded motif with four layers of base tetrads. Two of these are depicted in Figure 5c,d. Two additional layers resemble these and are related to them via the molecular twofold rotation axis. The above four layers are flanked by additional layers in which strands from two duplexes pair under formation of C−H···O hydrogen bonds between phosphate groups and C8(G) (Figure 5b).

Dimerization of homo-DNA duplexes in the crystal is stabilized by coordination of a magnesium hexahydrate ion that is located on the crystallographic dyad (Figure 4). The Mg^{2+} is bound to four phosphate groups from two duplexes near one of their termini. Mg^{2+} appears to be unique in its ability to establish these contacts as none of the other divalent metal ions tested resulted in crystal growth. Further lattice interactions involve stacking between terminal base pairs from symmetry-related duplexes.

Base−Backbone Inclination and Pairing Selectivity. Backbone torsion angles in homo-DNA fall into the *sc*$^-$, *ap*, *sc*$^+$, *sc*$^+$, *ap*, and *sc*$^-$ conformational ranges (α to ζ; Table 3). Deviations occur at the sites of bulged adenosines and in the region of T14 where the backbone exhibits high flexibility. T14 was modeled with both a boat and a chair conformation of the hexose (Figure 3a), and the former affects the conformation of some of the torsion angles of the neighboring T13 (ϵ) and C15 (α, *β*, *γ*) residues (Table 3). The sugar moiety of all other nucleotides is found in the expected energetically favorable chair conformation. Unlike the backbone angles, the χ torsion angles around glycosidic bonds exhibit considerable variation in the homo-DNA duplex (−62 to −126°, Table 3). These variations are evident from the irregular orientations of hexose sugars along strands (Figure 3a), a noticeable difference from the typically quite regular arrangement of 2′-deoxyriboses in A- and B-form

(23) Wang, A. H.-J.; Quigley, G. J.; Kolpak, F. J.; Crawford, J. L.; van Boom, J. H.; van der Marel, G.; Rich, A. *Nature* **1979**, *282*, 680−686.
(24) Han, G. W.; Kopka, M. L.; Langs, D.; Sawaya, M. R.; Dickerson, R. E. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 9214−9219.

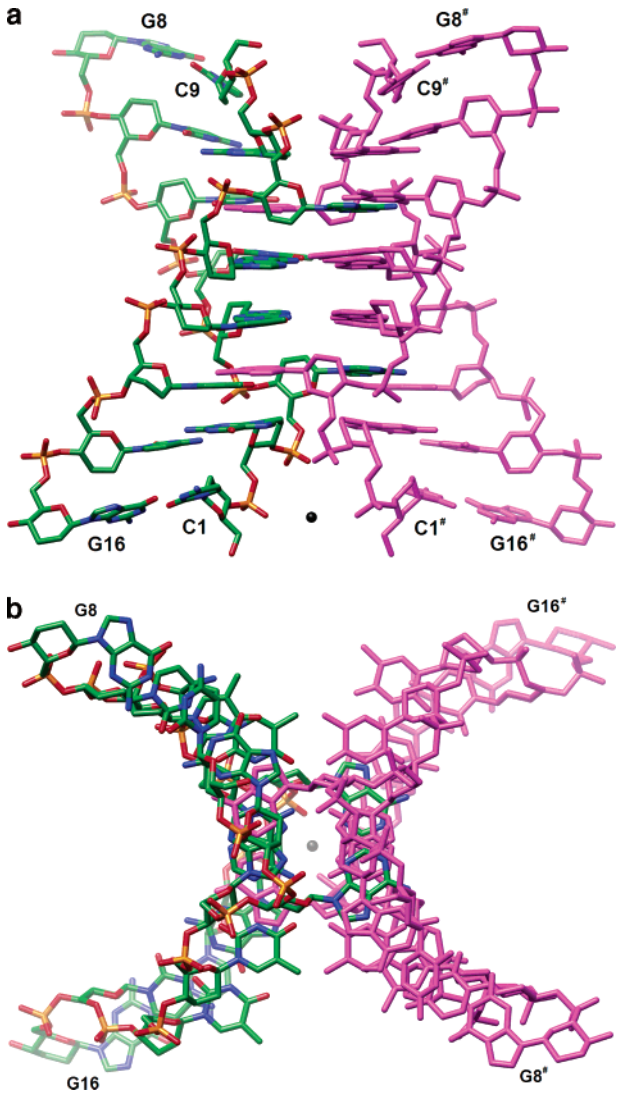

Figure 4. Dimerization of homo-DNA duplexes in the hexagonal crystal lattice. Two duplexes interacting via their convex surfaces viewed (a) perpendicular to the crystallographic twofold and (b) viewed along it. The color code for atoms of one duplex is the same as in Figure 3, and the symmetry-related duplex is shown with bonds in magenta. The Mg^{2+} ion is shown as a black sphere and marks the location of the dyad in (b).

DNA. The conformations of χ in A- and B-DNA duplexes are different, and the distribution of angles in the latter is much broader (Figure S2, Supporting Information). The average value of χ in homo-DNA is below those observed in A- (−160°) and B-DNA (−110°) and lies at the lower end of the angle range associated with B-form duplexes. To some extent, the variations in χ are the result of the extrusion of A3 and A11. However, as can be seen in the central ApT step, the hexose moieties of A4 (A12) and T5 (T13) are rotated relative to one another despite the virtual absence of a twist at that site (Figure 3b). The particular orientations of hexoses are most likely the result of the optimization of interstrand stacking, requiring subtle adjustments in χ and some of the backbone torsion angles (Table 3).

The relatively low twist in the homo-DNA duplex reveals a strong inclination of the backbones relative to the base-pair axes (Figure 3). The observed stacking between bases from opposite strands is a direct consequence of the pronounced backbone−base inclination. This parameter is not to be confused with the more familiar "inclination" parameter that describes the relative orientation of base pairs and either a local or the global helical axis.[25] The backbone−base inclination angle η_B is defined in Figure 6, and we have calculated η_B, defining the local backbone direction by either $P_n \rightarrow P_{n+1}$ vectors or a B-spline curve.[26] The angle determined in this fashion is independent of the helical twist (a more detailed description of the computer program for calculating η_B will be given elsewhere). The average backbone−base inclination in homo-DNA is 44° (Table 2). The inclination in B-form DNA is practically 0°, and in A-RNA it is ca. −30°. Therefore, the backbone−base inclination is related to the nature of the sugar in a nucleic acid-pairing system (obviously η_B can be calculated independently of whether an analogue is capable of pairing or not). One fundamental consequence of a large inclination is that the relative orientation of paired strands has to be antiparallel. Indeed, the homo-DNA pairing mode is strictly antiparallel,[11] and a parallel orientation of two base-paired RNA strands has never been observed. Conversely, DNA can form parallel-stranded duplexes.[27,28] Knowledge of the backbone−base inclination also allows a rationalization of the absence of cross-pairing between different nucleic acid systems or the existence thereof. Thus, homo-DNA does not pair with either DNA or RNA, an observation that can be understood on account of the differences in η_B that amount to ca. 45° and 75°, respectively.

Discussion

Although the homo-DNA octamer can be considered (using a somewhat arbitrary classification of molecular sizes) a small molecule (MW 2515 Da), it provides an excellent illustration of the challenges that can face the crystallographer when standard methods for determining a structure are seemingly exhausted. In the case of the homo-DNA crystal structure, the phosphoroselenoate (PSe) modification constitutes a sort of last resort approach. Provided that synthesis, purification, crystallization, and MAD data collection can all be carried out within about a week, the phasing power of a PSe derivative is excellent. The rate of oxidation to phosphate and the concomitant loss of selenium from crystals appear to depend on the packing density to some degree. Thus, crystals of a PSe derivative of left-handed Z-DNA that have an exceptionally low solvent content were stable for several weeks.[18] Compared to the use of brominated strands (Br^5U, Br^5C), PSe derivatization does not require synthesis of special (homo-DNA) nucleoside building blocks. Moreover, all phosphate groups can be targeted in principle, although in most cases the resulting pairs of diastereoisomeric PSe oligomers will need to be separated for producing derivative crystals. However, the synthesis of oligodeoxyribonucleotides with stereodefined internucleotide PSe functions has recently been described.[29]

The structure of $[dd(CGAATTCG)_2]$ analyzed here bears little resemblance to any of the previously described theoretical and experimental models of homo-DNA duplexes.[12,16,17] The most fundamental differences are a right-handed twist that exceeds 30° at two locations and a tighter than anticipated spacing of

(25) Dickerson, R. E. *Nucleic Acids Res.* **1989**, *17*, 1797−1803.
(26) Newman, W. N.; Sproull, R. F. *Principles of Interactive Computer Graphics*, 2nd ed.; McGraw-Hill: New York, 1979; Chapter 21.
(27) Germann, M. W.; Kalisch, B. W.; van de Sande, J. H. *Biochemistry* **1988**, *27*, 8302−8306.
(28) Otto, C.; Thomas, G. A.; Jovin, T. M.; Peticolas, W. L. *Biochemistry* **1991**, *30*, 3062−3069.
(29) Guga, P.; Maciaczek, A.; Stec, W. J. *Org. Lett.* **2005**, *7*, 3901−3904.

Table 2. Local Interbase Parameters,[a] Intrastrand P···P Distances, and Backbone−Base Inclination Angles (η_B)

base step	shift (d_x) (Å)	slide (d_y) (Å)	rise (d_z) (Å)	$P_n - P_{n+1}$ (Å)	tilt (deg)	roll (deg)	twist (deg)	η_B[b] (deg)
C1−G2	0.3	5.4	4.3		−3.8	12.8	1.0	57.6
G2−A3	10.7	8.3	3.2	5.5	−2.4	5.5	92.1	37.1
A3−A4	−8.7	−0.5	3.8	5.8	6.3	−5.2	−46.5	47.9
A4−T5	0.2	5.8	4.1	5.6	−5.9	5.6	10.4	45.1
T5−T6	0.3	2.0	3.2	6.2	4.3	−2.3	31.8	34.6
T6−C7	−0.3	3.6	3.3	5.8	11.0	10.1	9.3	43.1
C7−G8	0.1	6.1	4.6	6.1	−2.0	−4.0	0.1	
C9−G10	−0.6	6.0	5.1		25.0	21.1	2.2	47.7
G10−A11	−11.0	8.2	3.1	5.5	−1.7	1.4	94.7	34.5
A11−A12	8.6	0.0	4.1	5.8	−4.9	−6.5	−46.2	44.3
A12−T13	−0.4	5.8	4.2	5.5	7.9	5.3	8.1	47.4
T13−T14	1.0	3.7	2.9	6.1	−10.6	4.5	10.1	36.3
T14−C15	0.8	2.1	3.7	6.2	0.4	2.2	32.0	55.3
C15−G16	−0.4	5.4	4.1	5.3	3.0	17.7	1.4	
average	0.1	4.4	3.8	5.8	1.9	4.9	14.3	44.2

[a] Calculated with the program CURVES.[45] [b] Values refer to second base in step and were calculated using $P_n \rightarrow P_{n+1}$ vectors (Figure 6), omitting terminal bases.

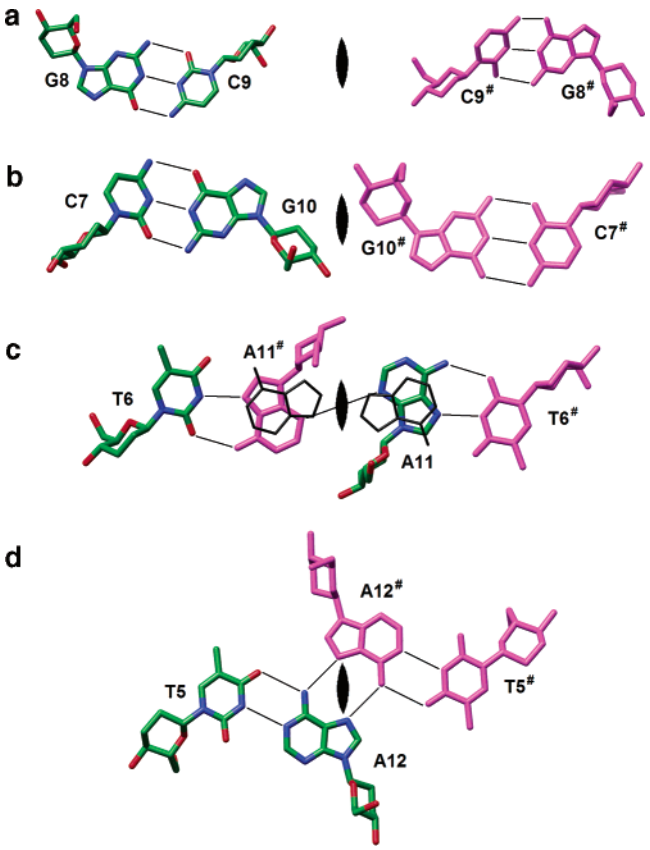

Figure 5. Intra- and interduplex base-pairing modes at individual levels of the homo-DNA octamer, starting with the bottom end. The position of the crystallographic twofold rotation axis is indicated. (a) G8−C9 and C9#−G8#. (b) C7−G10 and G10#−C7#. (c) T6−A11# and A11−T6#. A's in a standard Watson−Crick pairing mode opposite T6 and T6# are drawn with thin lines and superimposed on the observed reverse-Hoogsteen pairs to illustrate that the five-membered rings of adenines in the former would clash near the dyad. Extrusion of adenines prevents a clash, maintains two hydrogen bonds per A−T pair, possibly improves stacking, and leads to additional water-mediated hydrogen-bonding interactions between N3 atoms of looped-out A's. (d) T5−A12 and A12#−T5#. The relative arrangements of bases at the next levels in the paired duplexes are similar to those shown due to noncrystallographic twofold rotational symmetry. The color codes for atoms in the two duplexes are identical to those used in Figure 4, and hydrogen bonds are drawn with thin solid lines.

bases. In fact, at several steps the stacking distance is close to the ideal value of 3.4 Å. However, unlike in the familiar B- and A-form DNA duplexes, stacking in homo-DNA is almost exclusively of the interstrand type. Although this was anticipated on the basis of idealized linear models that essentially look like inclined ladders, duplexes without or with only very small twist also exhibited distances between adjacent base pairs that rendered impossible stabilizing stacking interactions. The crystal structure also reveals an irregular conformation of the duplex with values for twist and rise that vary greatly between individual base-pair steps (Figure 3a), notably at sites of extruded bases as a result of packing interactions involving pairs of duplexes. Remarkably, the central ApT step lacks a sizable twist and yet allows ideal stacking between A's from opposite strands (Figure 3b). The particular combination of backbone and glycosidic torsion angles and a strong slide between A−T base pairs prevents potential short contacts between hexose and base from adjacent intrastrand residues.

Homo-DNA constitutes an exception among hexopyranose-based oligonucleotide systems as oligomers with fully hydroxylated hexopyranose sugars, that is, the allo-, altro-, and glucopyranoses (Figure 1c−e) do not exhibit pairing (ref 2 and references cited therein). The structure of the homo-DNA duplex now allows a closer examination of the potential origins of the inability by these systems to pair. Hydroxyl groups attached to the 2′,3′-dideoxyglucopyranose of several residues in the various configurations display short contacts to either atoms from adjacent bases, sugars, or the phosphate group (Figure 7). Therefore, the crystal structure of homo-DNA points to overcrowding of the backbone as the most likely reason for the inability of fully hydroxylated hexopyranose-based nucleic acid analogues to form stable duplexes. Oligo-(4′ → 6′)-β-D-glucopyranosylnucleotides ("glucose nucleic acid") could not have served as a genetic coding system because the bulkiness of the sugar prevents orientations of bases that allow stacking interactions and Watson−Crick base pairing.

Linear models were studied for double-stranded DNA[30] and more recently have been considered for both homo-DNA[12,16] and pyranosyl-RNA.[31,32] Linear models of DNA show a distance of ca. 5 Å between adjacent intrastrand phosphorus atoms. Such

(30) Yagil, G.; Sussman, J. L. *EMBO J.* **1986**, *5*, 1719−1725.
(31) Pitsch, S.; Wendeborn, S.; Jaun, B.; Eschenmoser, A. *Helv. Chim. Acta* **1993**, *76*, 2161−2183.
(32) Schlönvogt, I.; Pitsch, S.; Lesueur, C.; Eschenmoser, A.; Jaun, B.; Wolf, R. M. *Helv. Chim. Acta* **1996**, *79*, 2316−2345.

Table 3. Summary of Backbone and Glycosidic Torsion Angles (All Values in Degrees)

nucleotide	α	β	γ	δ	ϵ	ζ	χ
C1			−163	68	−122	−54	−120
G2	−69	173	62	55	−165	−88	−64
A3[a]	−57	140	−174	72	−134	−84	−109
A4	−88	141	52	65	−152	−74	−81
T5	−57	174	60	55	136	−81	−65
T6[b]	−39	150	70	49	−168	−63	−125
C7	163	−157	−175	61	−124	−65	−117
G8	−75	179	74	57			−62
C9			58	62	−128	−59	−91
G10	−65	175	62	57	−167	−87	−62
A11[a]	−57	142	−177	65	−134	−81	−109
A12	−91	144	55	60	−150	−69	−83
T13	−59	174	65	53	166	−75	−70
T14[b]	−51	160	87	126[c]	−117	122	−88
C15	133	−92	137	55	−130	−58	−114
G16	−131	166	144	58			−92

[a] Looped-out of helix. [b] Pairs in a reverse-Hoogsteen mode with looped-out A. [c] Hexose in boat conformation; the torsion angles in the alternative backbone conformer with a chair conformation of the sugar are: −P(14)−α −75°−O6′−β 168°−C6′−γ 83°− C5′−δ 51°−C4′−ϵ −139°−O4′−ζ −42°−P(15)−α −109°−O6′−. The χ angle for residue T14 in the chair conformation is −122°.

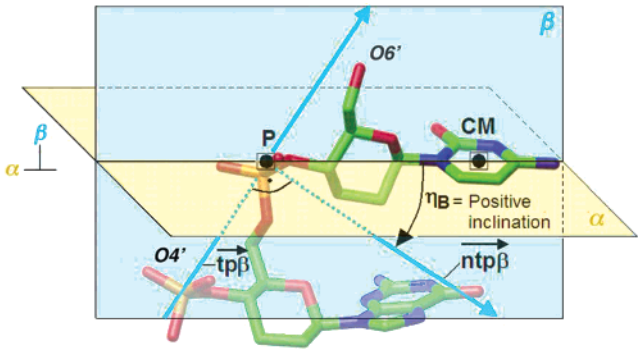

Figure 6. Backbone−base inclination η_B is defined as the angle between vectors P-CM and ntpβ, where (i) α is the best plane through the base$_n$, (ii) β is the plane normal to α and contains P (where the backbone, described by either the $P_n \rightarrow P_{n+1}$ vector or a B-spline curve through phosphorus atoms, pierces through α) as well as CM (the base's center of mass), (iii) tpβ is the orthogonal projection on plane β of tangent t to the backbone curve ($P_n \rightarrow P_{n+1}$ vector or B-spline curve) at position P, and (iv) ntpβ is the vector orthogonal to tpβ in plane β with P as origin. If tpβ is orthogonal to P-CM, the backbone is normal to the base plane and the inclination is 0° (i.e., canonical B-form DNA).

a spacing would lead to a severe electrostatic penalty, and without a stabilizing factor (i.e., a complex with a protein), repulsions between phosphate groups can only be relieved by twisting. Linear models of homo-DNA go along with P···P distances of >6 Å but lead to separation between adjacent base pairs that cannot account for effective stacking interactions. The crystal structure of the homo-DNA octamer duplex shows shorter P···P distances compared to that of the theoretical and NMR models, but the interactions do not appear to reach an electrostatically unfavorable range. A variable twist and strong sliding help avoid possible short contacts between hexose and base atoms from adjacent intrastrand residues. Distances between C2′ and C3′ atoms of hexoses and base atoms from 4′-adjacent residues (see Figures 1a and 3 for orientation) correspond to the sum of van der Waals radii in many cases.

Homo-DNA duplexes pair more strongly than DNA duplexes, and the higher stability of the former is entropy-based. The chair conformation of 2′,3′-dideoxyglucopyranose sugars observed in the homo-DNA crystal structure is in line with the favorable entropic contribution to pairing stability. Moreover, all α/ζ backbone torsion angle pairs fall into the *synclinal* range, illustrating the important influence of the anomeric effect in controlling the conformation of the phosphodiester moiety.[12] A possible backbone variant with both the α and γ torsion angles in the *antiperiplanar* range[12,16] is only present at residues C7 and C15 (Table 3). These residues lie adjacent to T's that pair with A's from a symmetry-related duplex in a reverse-Hoogsteen mode, and their conformation is most likely an indirect consequence of local geometrical changes in the duplex as a result of lattice interactions (Figure 4). The formation of reverse-Hoogsteen A−T pairs by antiparallel strands from interlocked duplexes (Figures 4a and 5c) is consistent with facile formation of purine−purine pairs by homo-DNA and the altered pairing priorities compared to DNA (G−G ≈ A−A > A−T).[11,13]

A striking property of homo-DNA that is not obscured by the modest twisting is the strong inclination between backbone and bases (Figure 3). This property dictates the observed prevalence of the interstrand base-stacking type and allows predictions regarding the relative polarity of paired strands in duplexes (antiparallel, ap, versus parallel, ps) and the potential for cross-pairing between different nucleic acid systems. It is straightforward to measure this parameter in a linear duplex, but we have developed a method for calculating the backbone−base inclination independent of the degree of twisting (Figure 6). Nucleic acid-pairing systems with significant positive or negative inclinations are essentially unable to form ps arrangements involving two or more oligonucleotides and reverse Watson−Crick base pairs. For example, RNA or homo-DNA cannot form ps duplexes or the so-called self-intercalated four-stranded DNA i-motif featuring hemiprotonated C−C$^+$ pairs.[33] Although such predictions are essentially qualitative, it appears that a difference of about 20° or more in the inclinations exhibited by two oligonucleotide systems will render cross-pairing impossible. DNA represents a unique system as its backbone−base inclination can vary between 0° (canonical B-form) and ca. −30° (canonical A-form), allowing it to self-pair in the ps and ap modes as well as cross-pair with RNA. Homo-DNA on the other hand is an autonomous pairing system as no oligonucleotide analogue that is able to cross-pair with homo-DNA has been identified to date.

(33) Gehring, K.; Leroy, J.-L.; Guéron, M. A. *Nature* **1993**, *363*, 561−565.

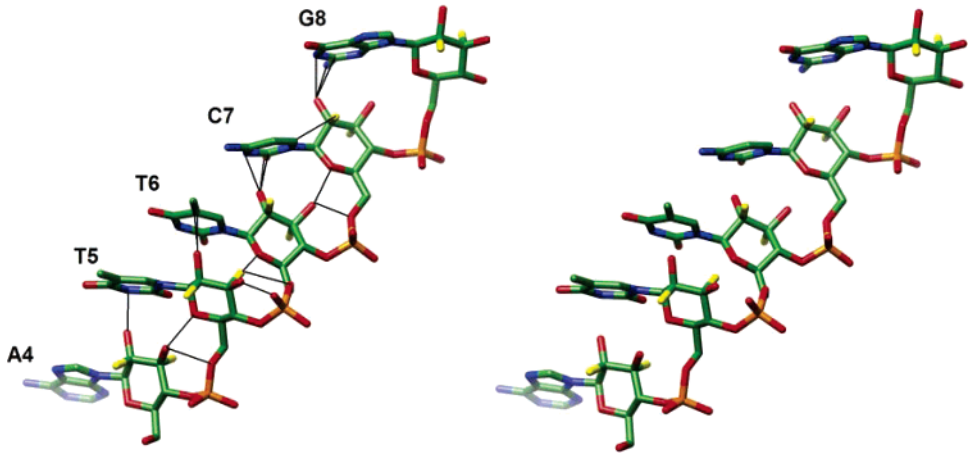

Figure 7. Insights from the homo-DNA structure into the absence of pairing with fully hydroxylated (4′ → 6′) hexopyranose nucleic acid systems. Stereo diagram depicting putative short contacts (thin solid lines; distances ranging from 1.8 to 2.8 Å) in the A4pT5pT6pC7pG8 stretch taken from the crystal structure between 2′- and 3′-hydroxyl groups of allo-, altro-, and glucopyranosylnucleotides (Figure 1c−e, respectively) and base, sugar, and/or phosphate moieties. Positions of hydroxyl oxygens (red for equatorial and yellow for axial positions) were calculated assuming a 1.4 Å C−O bond length and ideal tetrahedral geometry of hexose carbons.

The backbone−base inclination parameter observed here in the weakly wound homo-DNA duplex has less obvious implications for the pairing selectivity and stability with the natural nucleic acids that are related to its more evident role in affecting the relative contributions by intra- and interstrand stacking to duplex stability. For example, pRNA, which has an exceptionally strong negative inclination and is possibly more linear than homo-DNA, is capable of efficient nonenzymatic replication[34] (hampered in the case of RNA by self-pairing G-rich sequences) and self-assembly by ligative oligomerization.[35] RNA and pRNA cannot pair because of a considerable difference between their negative inclinations. However, a 2′-overhanging purine base in pRNA strongly enhances duplex stability, whereas a 4′-overhanging base does not[36] (Figure S3). Similarly, a 3′-overhanging purine base in RNA is stabilizing, whereas the effect of a 5′-overhanging base is negligible by comparison.[37,38] These differential stabilizations are a consequence of the negative inclination inherent to the RNA backbone. A biologically relevant example is found in transfer RNA where the 3′-overhanging nucleotides can be expected to exert a stabilizing effect.

The work presented here allows a structural rationalization for the inability of fully hydroxylated hexopyranosyl nucleic acids to form stable pairing systems. Thus, at least as far as informational base pairing, a fundamental and biologically indispensable property of the natural nucleic acids, is concerned, the structure of homo-DNA provides an answer to the question "why pentose and not hexose nucleic acids?". Naturally, this question can also be approached from another angle by focusing on the sugar moiety itself instead. For example, an experimental demonstration of the preferred formation of ribose and other pentoses from simple precursors compared to other types of sugars (i.e., hexoses) may point toward a scenario that could have precluded an evolutionary evaluation of hexose nucleic acids in terms of their ability to pair and self-replicate. Recent reports have shed light on this aspect (ref 39 and references cited therein), but the gathered evidence does not yet allow any definitive conclusions.

Materials and Methods

Preparation of Native, Phosphorothioate, and Phosphoroselenoate Homo-DNAs. The 2′,3′-dideoxyglucopyranose phosphoramidite G, A, C, and T building blocks and the 3′-terminal G controlled pore glass (CPG) support were prepared following described procedures.[10] All chemicals for solid-phase oligonucleotide synthesis were purchased from Glen Research (Sterling, VA). Reagents for PS homo-DNAs (3*H*-2,2-benzodithioiol-3-one)[40] and PSe homo-DNA (potassium selenocyanate) were purchased from Glen Research (Sterling, VA) and Aldrich (Milwaukee, WI), respectively. A 1% solution of 3*H*-2,2-benzodithioiol-3-one in acetonitrile was prepared prior to use. Saturated solutions of KSeCN in 95% acetonitrile/5% triethylamine were prepared by heating the mixture for 12 h and allowing it to cool to room temperature. The native homo-DNA octamer was synthesized following standard phosphoramidite protocols using CPG supports and solid-phase synthesizers (Pharmacia Gene Assembler or Applied Biosystems, Inc. 381A). The detritylated strand was deprotected using conc. NH_4OH (65 °C, 8 h) and HPLC-purified (RP-C4 column Rainin-Dynamax, 0.1 M TEAA pH 7.0, acetonitrile gradient). Homo-DNA octamers containing single phosphorothioate moieties were prepared according to the standard methods.[40] The homo-DNA phosphoroselenoate $CGA_{PSe}ATTCG$ was synthesized following a protocol described for synthesis of PSe-DNAs.[18] Diastereoisomers of the PS and PSe homo-DNA octamers were separated by strong anion exchange (SAX) HPLC using a DIONEX DNAPAC PA-100 analytical column (4 × 25 mm) purchased from Dionex Corp (Sunnyvale, CA). For preparatory runs, 10 OD units were purified at a time on an analytical column using a gradient of 25 mM TrisHCl (pH 7.8) to 0.5 M NaCl over 45 min with a flow rate of 1.0 mL/min (see Supporting Information for a chromatogram of the separation of the two diastereoisomeric PSe homo-DNA octamers). All oligonucleotides purified by HPLC were desalted on Sep-Pak cartridges (Waters, Inc.). Following desalting, oligonucleotide solutions were microfiltered and the concentration of stock solutions was adjusted to ca. 10 mM. Molecular weights of all oligonucleotides were determined by MALDI-TOF mass spectrometry.

(34) Pitsch, S.; Krishnamurthy, R.; Bolli, M.; Wendeborn, S.; Holzner, A.; Minton, M.; Lesueur, C.; Schlönvogt, I.; Jaun, B.; Eschenmoser, A. *Helv. Chim. Acta* **1995**, *78*, 1621−1635.
(35) Bolli, M.; Micura, R.; Eschenmoser, A. *Chem. Biol.* **1997**, *4*, 309−320.
(36) Micura, R.; Bolli, M.; Windhab, N.; Eschenmoser, A. *Angew. Chem., Int. Ed. Engl.* **1997**, *36*, 870−873.
(37) Petersheim, M.; Turner, D. H. *Biochemistry* **1983**, *22*, 256−263.
(38) Freier, S. M.; Burger, B. J.; Alkema, D.; Neilson, T.; Turner, D. H. *Biochemistry* **1983**, *22*, 6198−6206.
(39) Oberhuber, M.; Joyce, G. F. *Angew. Chem., Int. Ed.* **2005**, *44*, 7580−7583.
(40) Iyer, R. P.; Egan, W.; Regahn, J. B.; Beaucage, S. *J. Am. Chem. Soc.* **1990**, *112*, 1253−1254.

Crystallization and X-ray Data Collection. Crystallizations were performed at room temperature using the sitting-drop vapor diffusion method. Droplets containing either 1.5 mM native dd(CGAATTCG) octamer or the PS or PSe octamers, 6.25 mM sodium cacodylate, pH 7.0, 4 mM magnesium chloride, and 12% (v/v) MPD were equilibrated against a reservoir of 30% MPD. Crystals were mounted in nylon loops and directly frozen in liquid nitrogen. All diffraction data were collected on the 5-ID beamline of the DuPont-Northwestern-Dow Collaborative Access Team at the Advanced Photon Source (Argonne, IL) using various MARCCD detectors. MAD data were collected from a single crystal of the octamer dd(CGA$_{PSe}$ATTCG), and the precise wavelength of the selenium absorption edge was determined using a fluorescence detector. Data were collected at three wavelengths, using separate scans for high- and low-resolution reflections (Table 1). Numerous datasets were collected for native crystals over the years to optimize the resolution. The best data obtained to date have a resolution of 1.75 Å and are virtually 100% complete and of excellent quality (R_{merge} = 3.5%; Table 1). High-, medium-, and low-resolution frames were collected separately, and a final sweep with short exposure and attenuation was conducted to avoid overloads. All data were integrated and merged with HKL-2000.[20] Selected crystal data and data collection parameters for the PSe derivative and native crystals are summarized in Table 1.

Structure Determination and Refinement. MAD phasing was carried out with the program CNS[19] using data up to 2.1 Å resolution (Table 1). The initial model of the duplex was built by manually placing 2′-deoxyribonucleotides into the experimental map that was displayed with the program Turbo Frodo.[41] The structure was refined in CNS using standard DNA topology and parameter files. After numerous positional and B-factor refinement cycles as well as occasional rounds of annealing and manual readjustments, the all-DNA model resulted in values for R_{work} and R_{free} of 0.34 and 0.37, respectively, including all data between 30 and 2.1 Å resolution. At this point, the model was used for a few rigid body refinements with data collected from a native crystal. After additional refinement using data to 1.75 Å resolution in CNS, 2′,3′-dideoxy sugars with chair conformation were built into the electron density map and adaptations in the backbone torsion angles were made to switch from a model with pentose sugars to one with hexoses. Refinement was continued with the program REFMAC[21] as part of the *CCP*4 suite of programs.[42] Water molecules were placed into regions of overlaid ($2F_o - F_c$) sum and ($F_o - F_c$) difference electron density, and all homo-DNA atoms were treated with anisotropic temperature factors.[43,44] The current model has an *R*-factor of 0.23 for all data to 1.75 Å resolution. Selected refinement parameters are listed in Table 1. A detailed account of the efforts that led to the determination of the crystal structure of homo-DNA is provided elsewhere.[46]

Accession Codes. Final coordinates and structure factors have been deposited in the Protein Data Bank, http://www.rcsb.org (PDB ID 2H9S).

Acknowledgment. This article is dedicated to Professor Vladimir Prelog on the occasion of the centennial of his birthday. We are grateful for financial support by the U.S. National Institutes of Health (Grant R01 GM55237 to M.E.) and the Petroleum Research Fund administered by the American Chemical Society (Grant AC-34538 to M.E.). C.J.W. was the recipient of postdoctoral scholarships from the Natural Science and Engineering Research Council of Canada (NSERC) and the Fonds pour la Formation de Chercheurs et l'Aide a la Recherche (FCAR). We thank Dr. Zdzislaw Wawrzak for help with data collection and processing and Dr. Garib Murshudov for assistance with adaptation of the REFMAC parameter files. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science, under Contract No. W-31-109-Eng-38. The DuPont−Northwestern−Dow Collaborative Access Team (DND-CAT) Synchrotron Research Center at the Advanced Photon Source (Sector 5) is supported by E. I. DuPont de Nemours & Co., The Dow Chemical Company, the National Science Foundation, and the State of Illinois.

Supporting Information Available: Separation of homo-DNA phosphoroselenoate diastereoisomers, conformations of the glycosidic torsion angle in right-handed nucleic acid duplexes, and stabilizing effect of overhanging bases in nucleic acid duplexes. This material is available free of charge via the Internet at http://pubs.acs.org.

JA062548X

(41) Cambillau C.; Roussel A. *Turbo Frodo*, version OpenGL.1; Université Aix-Marseille II: Marseille, France, 1997.
(42) The *CCP*4 suite: Programs for protein crystallography. Collaborative computing project, number 4. *Acta Crystallogr., Sect. D* **1994**, *50*, 760−763.
(43) Murshudov, G. N.; Vagin, A. A.; Lebedev, A.; Wilson, K. S.; Dodson, E. J. *Acta Crystallogr., Sect. D* **1999**, *55*, 247−255.
(44) Winn, M. D.; Isupov, M. N.; Murshudov, G. N. *Acta Crystallogr., Sect. D* **2001**, *57*, 122−133.
(45) Lavery, R.; Sklenar, H. *J. Biomol. Struct. Dyn.* **1989**, *6*, 655−667.
(46) Egli, M.; Lubini, P.; Pallan, P. S. *Chem. Soc. Rev.* **2006**, *35*, in press.