J|A|C|S
ARTICLES
Published on Web 10/26/2002

# Crystal Structure of a B-Form DNA Duplex Containing (L)-α-Threofuranosyl (3′→2′) Nucleosides: A Four-Carbon Sugar Is Easily Accommodated into the Backbone of DNA

Christopher J. Wilds,[†] Zdzislaw Wawrzak,[‡] Ramanarayanan Krishnamurthy,[§] Albert Eschenmoser,[§,||] and Martin Egli*,[†]

*Contribution from the Department of Biological Sciences, Vanderbilt University, Nashville, Tennessee 37235, Argonne National Laboratory, DND-CAT Synchrotron Research Center, Sector 5, Advanced Photon Source, Argonne, Illinois 60439, The Skaggs Institute for Chemical Biology, The Scripps Research Institute, LaJolla, California 92037, and Laboratorium für Organische Chemie, Eidgenössische Technische Hochschule, CH-8093 Zürich, Switzerland*

Received June 4, 2002

***Abstract:*** (L)-α-Threofuranosyl-(3′→2′)-oligonucleotides (TNA) containing vicinally connected phosphodiester linkages undergo informational base pairing in an antiparallel strand orientation and are capable of cross-pairing with RNA and DNA. TNA is derived from a sugar containing only four carbon atoms and is one of the simplest potentially natural nucleic acid alternatives investigated thus far in the context of a chemical etiology of nucleic acid structure. Compared to DNA and RNA that contain six covalent bonds per repeating nucleotide unit, TNA contains only five. We have determined the atomic-resolution crystal structure of the B-form DNA duplex [d(CGCGAA)T*d(TCGCG)]$_2$ containing a single (L)-α-threofuranosyl thymine (T*) per strand. In the modified duplex base stacking interactions are practically unchanged relative to the reference DNA structure. The orientations of the backbone at the TNA incorporation sites are slightly altered in order to accommodate fewer atoms and covalent bonds. The conformation of the threose is C4′-*exo* with the 2′- and 3′-substituents assuming quasi-diaxial orientation.

## Introduction

To approach an understanding on a chemical level why ribo- and deoxyribonucleic acids, as opposed to other sugar-based nucleic acid systems, evolved as the genetic material is the goal of experimental studies toward a chemical etiology of nucleic acid structure.[1] Systematic screening of the base-pairing properties of nucleic acids with alternative sugars, including hexopyranoses, pentopyranoses, and tetrofuranoses, has revealed that the capacity for Watson−Crick base pairing is widespread among potentially natural sugar-based nucleic acid analogues. Nucleic acid analogues that lack this fundamental property can be ruled out as candidates that could have been selected as RNA alternatives, whereas analogues that have similar or superior capacity for base pairing merit further investigation.

One of the simplest nucleic acid analogues found to base pair is the one based on the sugar (L)-α-threose, an aldo sugar that contains only four carbon atoms.[2] (L)-α-Threofuranosyl oligonucleotides (TNA) with a 3′→2′ linkage (Figure 1) exhibit efficient base pairing and were shown to cross-pair with both RNA and DNA. Recently, TNAs with either 3′- or 2′-phosphoramidate linkages were also found to base pair in the Watson−Crick mode and to undergo cross-pairing with TNA, RNA, and DNA.[3] Moreover, replacement of adenine by 2,6-diaminopurine considerably enhances the stability of TNA:TNA, TNA:RNA, and TNA:DNA duplexes and accelerates template-directed ligation of TNA ligands.[4]

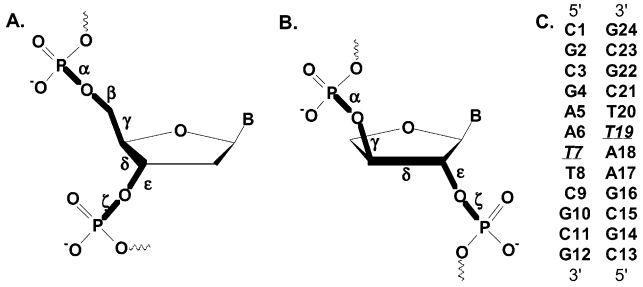


***Figure 1.*** Topology and torsion angles for (A) DNA and (B) TNA. The internucleotide linkages are shown in bold illustrating the different connectivities and number of covalent bonds per repeating nucleotide unit (six for DNA and five for TNA). (C) The sites of TNA incorporation in the Dickerson−Drew dodecamer at positions T7 and T19.

A significant difference between the (L)-α-threofuranosyl (3′→2′) and pentofuranosyl (5′→3′) backbone linkages is the

* To whom correspondence should be addressed. E-mail: martin.egli@vanderbilt.edu.
[†] Vanderbilt University.
[‡] DND-CAT Synchrotron Research Center.
[§] The Scripps Research Institute.
[||] Laboratorium für Organische Chemie.

(1) Eschenmoser, A. *Science* **1999**, *284*, 2118−2124.
(2) Schöning, K. U.; Scholz, P.; Guntha, S.; Wu, X.; Krishnamurthy, R.; Eschenmoser, A. *Science* **2000**, *290*, 1347−1351.
(3) Wu, X.; Guntha, S.; Ferencic, M.; Krishnamurthy, R.; Eschenmoser, A. *Org. Lett.* **2002**, *4*, 1279−1282.
(4) Wu, X.; Delgado, G.; Krishnamurthy, R.; Eschenmoser, A. *Org. Lett.* **2002**, *4*, 1283−1286.


**10.1021/ja0207807 CCC: $22.00** 

number of covalent bonds between consecutive phosphates: five as opposed to six, respectively (Figure 1). To better understand the conformational properties of TNA that allow it to pair with DNA, we replaced a single 2′-deoxy-T (T) in the Dickerson−Drew dodecamer (DDD) CGCGAAT*TCGCG with (L)-α-threofuranosyl-T (T*) and determined the crystal structure of the modified B-form duplex at 1.2 Å resolution. Here, we analyze the structural consequences of the DNA (T) → TNA (T*) replacement for the duplex geometry and describe the conformational details of the TNA units in the B-form DNA environment.

## Results and Discussion

**X-ray Crystallography.** One of the surprising outcomes of the analysis of the pairing properties of TNA, particularly considering the reduced number of atoms in the TNA backbone, is its efficient cross-pairing with both DNA and RNA. This observation implies that TNA can adapt to the geometric constraints imposed by the A- and B-form double helices much like DNA itself (alternatively, one could argue that DNA can adapt to the geometric constraints imposed by TNA). In the present investigation we wished to obtain insight into the geometric properties of TNA that allow it to pair with DNA under formation of a B-form duplex. The oligodeoxynucleotide with the sequence CGCGAATTCGCG, the DDD, was chosen as the crystallization template. We previously used this dodecamer to analyze the conformations of antisense nucleic acid analogues.[5−9] The native dodecamer and three modified DDDs with either one or both Ts replaced by T*, CGCGAAT*TCGCG, CGCGAATT*CGCG, and CGCGAAT*T*CGCG, were synthesized by solid-phase methods following standard protocols.

The three modified DDDs were subjected to hanging and sitting drop crystallization trials using commercially available sparse matrix screens (Hampton Research, Laguna Niguel, CA).[10] The best crystals were obtained for the duplex with TNA residues at positions T7 and T19 (Figure 1). Prior to the structural analysis circular dichroism (CD) spectra were recorded for the native DDD and the modified duplexes. These spectra indicated close resemblance between the native duplex and those containing TNA residues (Figure 2). Diffraction data to 1.20 Å resolution for the above DDD with single TNA nucleosides per strand were collected at the Advanced Photon Source synchrotron. The unit cell dimensions and symmetry of the TNA-modified DDD crystal are indicative of a lattice that is isomorphous with that of the native duplex (Table 1).

Rotation and translation searches with an all-DNA DDD[11] confirmed the identical orientation of the modified duplex in the orthorhombic unit cell. For the initial refinement of the model with the programs CNS[12] and SHELXL-97[13] all residues were treated as 2′-deoxyribonucleosides. Only during the final

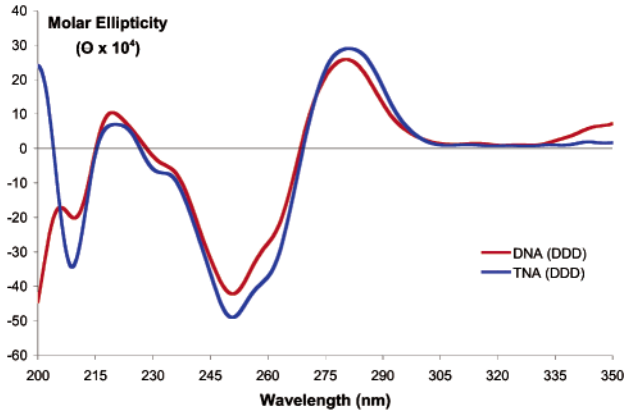


***Figure 2.*** CD spectra of the native (red) and TNA-modified (blue) DDD duplexes. The spectra were obtained at 5 °C and are the average of five scans.

***Table 1.*** Selected Crystal Data and Refinement Parameters

| | |
|---|---|
| space group | orthorhombic $P2_12_12_1$ |
| cell constants (Å) | $a = 25.28$, $b = 40.08$, $c = 65.86$ |
| resolution (Å) | 1.20 |
| no. of unique reflections (20−1.20 Å) | 21 311 |
| completeness (1.24−1.20 Å) (%) | 99.8 (98.2) |
| $R_{merge}$ (1.24 −1.20) (%) | 3.7 (32.7) |
| $R$-factor (%, no $\sigma$ cutoff) | 16.0 |
| no. of refinement parameters | 5791 |
| data-to-parameters ratio | 3.9 |
| no. of restraints | 13 838 |
| no. of DNA atoms (incl. H's) | 753 |
| no. of waters | 163 |
| no. of ions | 1 $Mg^{2+}$ |
| rms bonds (Å) | 0.007 |
| rms angles (1···3 dist., Å) | 0.014 |

stages of the anisotropic refinement and following the addition of a considerable number of water molecules were the connectivities of the T7 and T19 residues altered so as to reflect the constitution of the TNA backbone. This was possible because the locations of the sugar and base portions of the (L)-α-threofuranosyl residues are quite similar to those of the corresponding 2′-deoxyribo Ts (Figure 3). Consequently, the parameter and topology files were updated to accurately represent the bond lengths and angles of the threose. The values for individual parameters were taken from the crystal structure of the (L)-α-threofuranosyl-T nucleoside.[2] Subsequently, the structure was further refined, more water molecules were added, and the positions of DNA hydrogen atoms were calculated based on those of covalently bound heavy atoms and included in the refinement. The final $R$-factor is 16.0% for all 21 311 reflections to 1.20 Å resolution (Table 1). Examples of the final ($2F_o - F_c$) Fourier sum electron density are shown in Figures 3 and 4. Calculated helical parameters and backbone torsion angles are listed in Tables 2 and 3, respectively.

**Overall Conformation of the TNA-Modified DDD.** As expected from the similarities between the CD spectra (Figure 2) of the native and TNA-modified DDDs as well as the virtually identical unit cell constants for crystals of the two duplexes (Table 1), their overall conformations vary only minimally

(5) Portmann, S.; Altmann, K.-H.; Reynes, N.; Egli, M. *J. Am. Chem. Soc.* **1997**, *119*, 2396−2403.
(6) Tereshko, V.; Gryaznov, S.; Egli, M. *J. Am. Chem. Soc.* **1998**, *120*, 269−283.
(7) Egli, M. In *Advances in Enzyme Regulation*; Weber, G., Ed.; Elsevier Science Ltd.: Oxford, U.K., 1998; Vol. 38, pp 181−203.
(8) Berger, I.; Tereshko, V.; Ikeda, H.; Marquez, V. E.; Egli, M. *Nucleic Acids Res.* **1998**, *26*, 2473−2480.
(9) Teplova, M.; Minasov, G.; Tereshko, V.; Inamati, G.; Cook, P. D.; Manoharan, M.; Egli, M. *Nat. Struct. Biol.* **1999**, *6*, 535−539.
(10) Berger, I.; Kang, C.; Sinha, N.; Wolters, M.; Rich, A. *Acta Crystallogr.* **1996**, *D52*, 465−468.
(11) Shui, X.; McFail-Isom, L.; Hu, G. G.; Williams, L. D. *Biochemisty* **1998**, *37*, 8341−8355.
(12) Brünger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gros, P. R.; Grosse-Kunstleve, W.; Jiang, J.-S.; Kuszewski, J.; Niges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, K. L. *Acta Crystallogr.* **1998**, *D54*, 905−921.
(13) Sheldrick, G. M.; Schneider, T. R. *Methods Enzymol.* **1997**, *277*, 319−343.

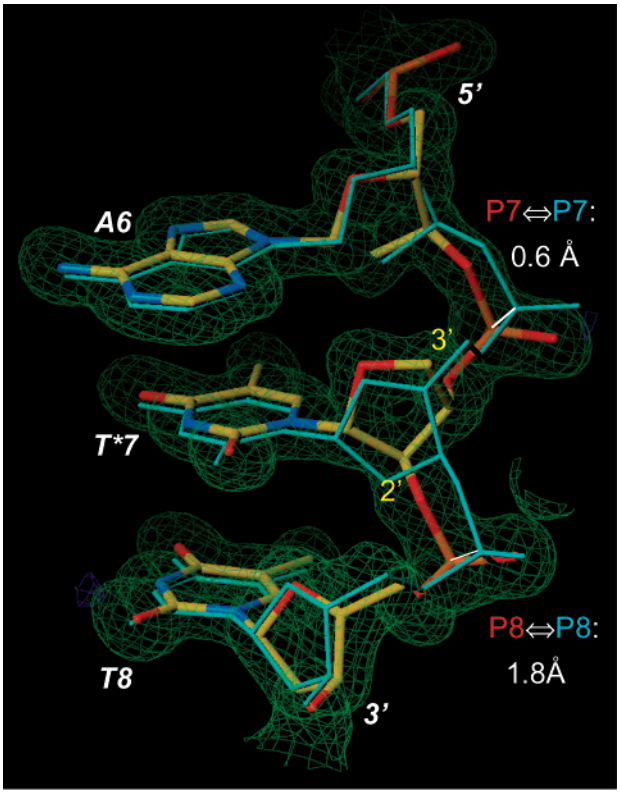


***Figure 3.*** Quality of the structure of the TNA-modified DDD. Final ($2F_o - F_c$) Fourier sum electron density map (1.4 $\sigma$ level) around residues A6, T*7 and T8. The DDD reference duplex (NDB Berman, H. M.; Zardecki, C.; Westbrook, J. *Acta Crystallogr*. **1998**, *D54*, 1095−1104.) code BDL084[11]) is drawn with thin blue lines. Atoms of the TNA-modified strand are colored yellow, red, blue, and orange for carbon, oxygen, nitrogen, and phosphorus, respectively. Residue T*7 is a TNA sugar unit with a C4′-*exo* conformation. Only base atoms of residues A6 and T8 were used to generate the superposition.

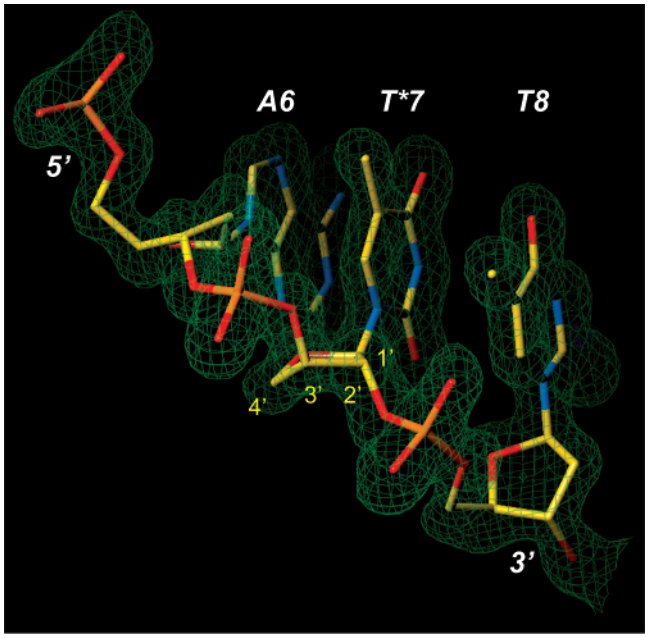


***Figure 4.*** The C4′-*exo* conformation of the L-(α)-threofuranose (center) in the modified DDD duplex [d(CGCGAA)T*d(TCGCG)]$_2$. Final ($2F_o - F_c$) Fourier sum electron density map (1.4$\sigma$ level) around residues A6, T*7, and T8 is shown as a green meshwork, and the color coding of nucleic acid atoms is identical to that in Figure 3.

(Figure 5). Selected helical parameters for the modified duplex and deviations from the reference duplex are summarized in Table 2. The root-mean-square (rms) deviation for atoms of the two duplexes based on a superposition of all 22 phosphorus atoms amounts to 0.41 Å. The biggest deviation occurs between the phosphate groups of T8 residues (1.64 Å) and is therefore adjacent to the TNA incorporation site (Figure 3). In both strands, significant deviations between the native and chemically modified duplexes are restricted to the locations of threose nucleosides and the 5′- and 3′-adjacent phosphate groups (Figures 3 and 5).

**Threose Sugar Conformation.** The threoses in both strands adopt C4′-*exo* pucker (Figure 4, Table 3), thus contributing to a considerable change in the local backbone geometry relative to the reference duplex (Figures 3 and 5). The C4′-*exo* and C3′-*endo* conformations occupy adjacent ranges in the furanose pseudorotation cycle.[14] In addition to the two TNA thymidine residues, the sugars of residues C3 and C15 also exhibit C4′-*exo* pucker (Table 3), a conformation that is rather unusual given the overall B-form geometry of the DDD duplex. In the case of the threose sugars, the particular pucker brings about a quasi-trans-diaxial orientation of the 3′ and 2′ substituents (Figure 4). The sugar conformation observed here in a B-DNA environment closely resembles that seen in the crystal structure of the (L)-α-threofuranosyl thymine nucleoside.[2] Such an arrangement helps maximize the distance separating the 3′ and 2′ phosphate groups. At first sight it may appear as if the reduction of backbone atoms per repeating unit from six to five in TNA relative to DNA may render the former incompatible with a standard B-form geometry (Figure 1). However, a quasi-diaxial orientation of the 3′ and 2′ substituents allows TNA to bridge a much wider gap than originally anticipated. The distance between P7* (an asterisk designates an atom belonging to a TNA residue) and P8 is 5.80 Å. In the second strand, the distance between P19* and P20 is 6.09 Å. Thus, the spacings between these phosphates are somewhat below the average distance of 6.64 Å between adjacent phosphates in the TNA-modified duplex. Remarkably, the reduced distance between neigboring phosphates and the different constitution of the tetrose sugar compared to the standard pentose causes conformational changes only at the local level and without affecting base stacking (Figures 3 and 5).

**TNA Backbone Conformation.** The backbone conformations of DNA and RNA containing 2′-deoxyribofuranose and ribofuranose, respectively, can be described with six torsion angles α to ζ (Figure 1A). In an analogous fashion, we have defined five torsion angles α, $\gamma$, $\delta$, $\epsilon$, and ζ to characterize the backbone conformations of the TNA portions in the modified DDD (Figure 1B). Since the bridging 3′-phosphate oxygen in TNA is directly attached to the five-membered ring (rather than being separated from it by a methylene carbon as is the case for DNA, Figure 1), it seemed reasonable to leave out the $\beta$ torsion angle for comparing the backbone geometries of TNA and DNA (Table 3). Accordingly, the torsion angle ranges for residue T*7 are *sc*-, *ap*, *ap*, *sc*-, and *ac*- (α to ζ). For residue T*19, they are *sc*-, *ap*, *ap*, *ac*-, and *sc*-. In standard B-form DNA, the six backbone torsion angles fall into the *sc*-, (*ap*), $sc^+$, *ap*, *ap*, and *sc*- ranges ($\beta$ in parentheses). While the orientations of individual bonds and the torsions around them clearly differ between TNA

(14) Saenger, W. In *Principles of Nucleic Acid Structure*; Springer-Verlag Inc.: New York, 1984.

***Table 2.*** Selected Helical Parameters and Deviations from the DDD Reference Duplex (in Parentheses)[a]

| (global base–base) | shear $S_x$ (Å) | stretch $S_y$ (Å) | stagger $S_z$ (Å) | buckle $\kappa$ (deg) | propeller $\omega$ (deg) | opening $\sigma$ (deg) |
|---|---|---|---|---|---|---|
| C(1)−G(24) | 0.2 (0.0) | 0.1 (−0.1) | −0.1 (0.1) | 5.0 (2.3) | −16.8 (−6.1) | 2.4 (−3.1) |
| G(2)−C(23) | −0.5 (0.5) | 0.0 (0.0) | 0.4 (0.0) | 9.3 (1.5) | −13.5 (−5.8) | 1.3 (−2.1) |
| C(3)−G(22) | 0.0 (0.1) | 0.0 (−0.1) | 0.2 (0.0) | −6.9 (0.9) | −3.0 (−5.7) | 1.8 (−2.1) |
| G(4)−C(21) | −0.1 (0.0) | 0.1 (0.0) | 0.0(0.0) | 14.7 (−1.7) | −7.3 (−4.0) | −0.6 (2.4) |
| A(5)−T(20) | 0.0 (0.0) | −0.1 (0.0) | −0.1 (0.1) | 11.7 (−5.8) | −14.1 (−5.8) | 4.8 (−3.2) |
| **A(6)−T*(19)** | **0.4 (−0.4)** | **0.0 (0.0)** | **−0.2 (0.2)** | **5.1 (−4.0)** | **−24.2 (5.3)** | **−0.9 (5.9)** |
| **T*(7)−A(18)** | **−0.2 (0.2)** | **−0.2 (0.1)** | **−0.1(0.2)** | **−0.5 (−0.6)** | **−19.8 (−0.4)** | **−0.9 (5.2)** |
| T(8)−A(17) | −0.1 (0.0) | −0.1 (0.0) | −0.2 (0.2) | −1.5 (0.0) | −8.5 (−14.6) | 2.1 (1.4) |
| C(9)−G(16) | −0.3 (0.3) | 0.1 (0.0) | 0.0 (0.0) | −13.7 (−0.8) | −9.5 (−3.6) | −1.3 (0.2) |
| G(10)−C(15) | 0.1 (0.1) | 0.1 (0.0) | 0.3 (−0.2) | 8.4 (−2.5) | −3.5 (−6.7) | 7.9 (−2.7) |
| C(11)−G(14) | 0.2 (0.0) | −0.1 (0.1) | −0.1 (0.2) | 0.5 (−2.4) | −19.6 (−3.4) | −5.6 (1.7) |
| G(12)−C(13) | 0.1 (−0.1) | 0.0 (0.1) | 0.1 (0.1) | 3.7 (1.7) | −7.5 (0.2) | −4.9 (2.7) |

| (local inter-base pair) | shift $D_x$ (Å) | slide $D_y$ (Å) | rise $D_z$ (Å) | tilt $\tau$ (deg) | roll $\rho$ (deg) | twist $\pi$ (deg) |
|---|---|---|---|---|---|---|
| C(1):G(24)/G(2):C(23) | 0.1 (0.0) | −0.3 (−0.1) | 3.2 (0.0) | −3.8 (0.5) | 5.2 (3.7) | 32.9 (1.7) |
| G(2):C(23)/C(3):G(22) | 0.5 (−0.1) | 0.0 (0.1) | 3.7 (0.0) | 2.3 (1.3) | −6.8 (−6.1) | 43.6 (−0.9) |
| C(3):G(22)/G(4):C(21) | −0.4 (0.3) | 0.2 (0.0) | 3.0 (0.0) | 1.6 (−0.4) | 7.7 (8.9) | 28.1 (−1.8) |
| G(4):C(21)/A(5):T(20) | −0.1 (−0.6) | −0.8 (0.1) | 3.4 (0.1) | 0.3 (−3.7) | 2.0 (−2.7) | 33.9 (3.0) |
| **A(5):T(20)/A(6):T*(19)** | **−0.1 (0.2)** | **−1.1 (0.3)** | **3.4 (−0.1)** | **−0.7 (0.2)** | **7.3 (−7.8)** | **38.4 (−1.2)** |
| **A(6):T*(19)/T*(7):A(18)** | **0.1 (0.0)** | **−0.7 (−0.3)** | **3.4 (−0.2)** | **−1.3 (0.9)** | **−9.9 (5.9)** | **32.0 (1.1)** |
| **T*(7):A(18)/T(8):A(17)** | **0.0 (−0.2)** | **−1.0 (0.1)** | **3.2 (0.0)** | **0.8 (0.4)** | **3.4 (−3.1)** | **35.5 (−1.7)** |
| T(8):A(17)/C(9):G(16) | −0.2 (0.1) | −1.0 (0.2) | 3.5 (0.0) | 0.8 (1.1) | −3.4 (−2.8) | 39.8 (1.2) |
| C(9):G(16)/G(10):C(15) | −1.0 (−0.1) | 0.5 (−0.1) | 2.9 (0.1) | −4.4 (1.2) | 2.7 (6.5) | 29.8 (−1.0) |
| G(10):C(15)/C(11):G(14) | −1.6 (0.1) | 0.0 (−0.1) | 3.4 (0.0) | −0.3 (−3.2) | −9.7 (−4.2) | 41.8 (−1.5) |
| C(11):G(14)/G(12):C(13) | −0.1 (−0.1) | −0.2 (−0.1) | 3.3 (−0.1) | −0.3 (0.9) | 5.8 (4.2) | 35.1 (0.0) |

[a] Calculated with the program Curves. (a) Lavery, R.; Sklenar, J. *J. Biomol. Struct. Dyn.* **1988**, *6*, 63−91. (b) Lavery, R.; Sklenar, H. *J. Biomol. Struct. Dyn.* **1989**, *7*, 655−667.

***Table 3.*** Backbone Torsion Angles, Glycosyl Angles, and Sugar Puckers (Deviations from DDD Reference Structure in Parentheses)[a]

| residue | α (deg) | β (deg) | γ (deg) | δ (deg) | ϵ (deg) | ξ (deg) | χ (deg) | *P* (deg) | amplitude | pucker |
|---|---|---|---|---|---|---|---|---|---|---|
| C(1) | | | −64 (−6) | 142 (2) | −168 (−4) | −96 (−2) | −104 (−2) | 170 (−6) | 35 (1) | C2′-*endo* |
| G(2) | −68 (−2) | −172 (0) | 44 (−1) | 133 (15) | −164 (12) | −126 (−31) | −98 (13) | 142 (18) | 39 (−1) | C1′-*exo* |
| C(3) | −54 (15) | 150 (−19) | 52 (−2) | 94 (−1) | −171 (6) | −86 (4) | −136 (3) | 59 (6) | 33 (1) | C4′-*exo* |
| G(4) | −62 (−3) | −178 (−8) | 52 (−2) | 140 (5) | 178 (15) | −110 (−35) | −98 (4) | 160 (−4) | 34 (4) | C2′-*endo* |
| A(5) | −57 (9) | 173 (−17) | 52 (2) | 134 (−1) | 178 (4) | −90 (−1) | −112 (−7) | 147 (12) | 35 (−3) | C2′-*endo* |
| A(6) | −70 (5) | −175 (2) | 45 (1) | 134 (−7) | −180 (−3) | −96 (0) | −103 (−6) | 145 (8) | 41 (−10) | C2′-*endo* |
| **T*(7)** | **−84 (35)** | | **−173 (222)** | **157 (−44)** | **−100 (−77)** | **−136 (40)** | **−104 (−16)** | **54 (72)** | **41 (−6)** | **C4′-*exo*** |
| T(8) | **−47 (−7)** | **150 (18)** | 53 (1) | 137 (−23) | 174 (−3) | −92 (−3) | −121 (1) | 156 (−30) | 34 (3) | C2′-*endo* |
| C(9) | −63 (10) | −168 (−5) | 51 (0) | 132 (6) | −164 (8) | −86 (−10) | −114 (2) | 150 (8) | 34 (−1) | C2′-*endo* |
| G(10) | −69 (8) | 169 (−6) | 42 (−3) | 140 (3) | −98 (−1) | 147 (0) | −90 (7) | 143 (3) | 50 (−5) | C1′-*exo* |
| C(11) | −71 (−2) | 146 (−2) | 50 (0) | 138 (5) | −164 (0) | −86 (−40) | −112 (−1) | 154 (10) | 41 (−3) | C2′-*endo* |
| G(12) | −69 (122) | −179 (−37) | 27 (−93) | 114 (34) | | | −99 (20) | 110 (97) | 26 (0) | C1′-*exo* |
| C(13) | | | 50 (8) | 137 (2) | −166 (−3) | −95 (−10) | −108 (2) | 167 (−3) | 28 (3) | C2′-*endo* |
| G(14) | −57 (−1) | 170 (−4) | 41 (8) | 114 (1) | 175 (3) | −90 (−4) | −114 (3) | 110 (20) | 34 (−1) | C1′-*exo* |
| C(15) | −58 (1) | 164 (−4) | 55 (2) | 86 (−4) | −173 (−1) | −85 (0) | −142 (4) | 41 (0) | 40 (−2) | C4′-*exo* |
| G(16) | −59 (−3) | 176 (−1) | 58 (4) | 147 (−6) | 157 (−6) | −94 (4) | −103 (1) | 171 (−2) | 36 (−4) | C2′-*endo* |
| A(17) | −61 (5) | −168 (7) | 57 (−3) | 143 (3) | −179 (1) | −84 (−7) | −112 (4) | 172 (1) | 36 (−4) | C2′-*endo* |
| A(18) | −74 (13) | 179 (−4) | 46 (0) | 126 (−13) | 179 (−13) | −88 (−3) | −110 (−3) | 130 (−2) | 43 (−10) | C1′-*exo* |
| **T*(19)** | **−86 (37)** | | **−170 (222)** | **160 (−32)** | **−130 (−43)** | **−98 (−9)** | **−103 (−13)** | **45 (98)** | **46 (−11)** | **C4′-*exo*** |
| T(20) | **−59 (14)** | **−179 (−9)** | 54 (−10) | 136 (−1) | −176 (12) | −85 (−21) | −113 (4) | 156 (−4) | 32 (6) | C2′-*endo* |
| C(21) | −69 (11) | 175 (−16) | 49 (3) | 103 (−7) | −172 (2) | −85 (4) | −124 (0) | 100 (−7) | 41 (0) | O4′-*endo* |
| G(22) | −64 (−3) | 175 (2) | 43 (2) | 142 (1) | −149 (0) | −166 (−7) | −93 (7) | 148 (4) | 45 (−4) | C2′-*endo* |
| C(23) | −59 (2) | 138 (−9) | 49 (3) | 88 (−4) | −163 (2) | −77 (0) | −149 (−1) | 20 (−3) | 41 (−4) | C3′-*endo* |
| G(24) | −68 (6) | 167 (2) | 59 (−6) | 81 (6) | | | −141 (−1) | 11 (2) | 40 (−5) | C3′-*endo* |

[a] Backbone torsion angles defined as: O2′-P-α-O3′-γ-C3′-δ-C2′-ϵ-O2′-ξ-P′-O3′ (TNA); O3′-P-α-O5′-β-C5′-γ-C4′-δ-C3′-ϵ-O3′-ξ-P−O5′ (DNA).

and DNA, it is noteworthy that the backbone torsion angles immediately before and after the TNA residues exhibit no significant deviations from those in the remaining DNA portions in the modified DDD duplex. Thus, the ζ torsion angles of both residue A6 and A18 preceding a TNA nucleotide fall into the *sc*- range, and the actual values differ only minimally from those of the corresponding angles in the reference DDD (Table 3). Similarly, the α and β torsion angles of residues T8 and T20 following a TNA residue fall into the standard *sc*- and *ap*, ranges, respectively. The values of all four angles in the TNA-modified duplex differ only marginally from those in the native DDD (Table 3).

Figures 3 and 6 illustrate nicely that the conformational perturbations caused by incorporation of TNA into B-form DNA are basically limited to the TNA nucleotide itself as well as the 2′-phosphate group. In the case of T*7, the shift of the 3′-phosphorus relative to the corresponding P in the reference DDD is 0.6 Å and for the 2′-phosphorus the shift amounts to 1.8 Å (Figure 3). The shifts of the equivalent 3′-P and 2′-P atoms for residue T*19 on the opposite strand relative to the native duplex are 1.1 and 0.3 Å, respectively. These displacements have an axial and a radial component, with the TNA backbone shifted toward the 3′-end of the DDD strand (Figure 6A) and

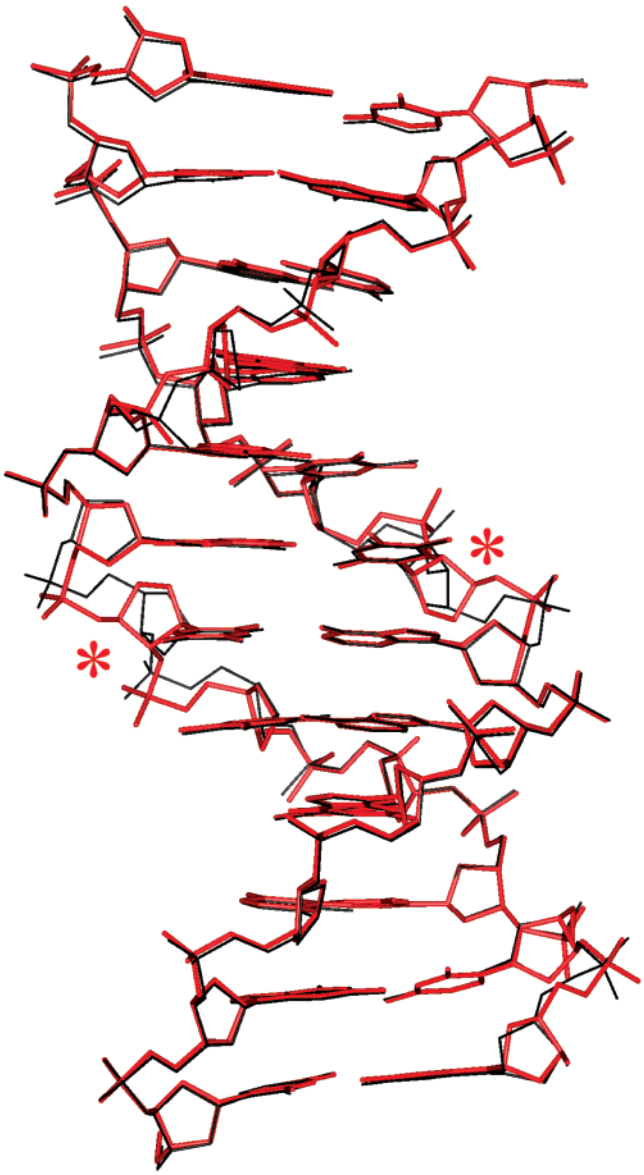

***Figure 5.*** Superposition (using positions of phosphorus atoms) of the TNA-modified (red) and native[11] (black) DDD duplexes. The C1:G24 base pair is at the top, and asterisks denote the locations of TNA residues.

toward the helix axis (Figure 6B). Therefore, an all-TNA duplex will have a slightly smaller diameter than B-form DNA (Figure 6B).

Consistent with the thermodynamic stability of TNA duplexes and the ability of TNA to pair stably with both DNA and RNA,[2] the stacking interactions in the center of the modified duplex are virtually unchanged compared to the reference DDD. As visible in Figure 6B, the orientations of the glycosyl bonds of residues A6, T7/T7*, and T8 are very similar in the two duplexes, and the locations of the nucleobases are nearly identical. As in duplex DNA, TNA nucleosides adopt anti conformations around the glycosyl bond (Table 3). Another feature that is more or less unaffected by the incorporation of alternative sugars into the DDD backbone here is the "inner" spine of hydration in the minor groove.[15] Since the positions of O4′ atoms in the tetrose sugars are very similar to those of the

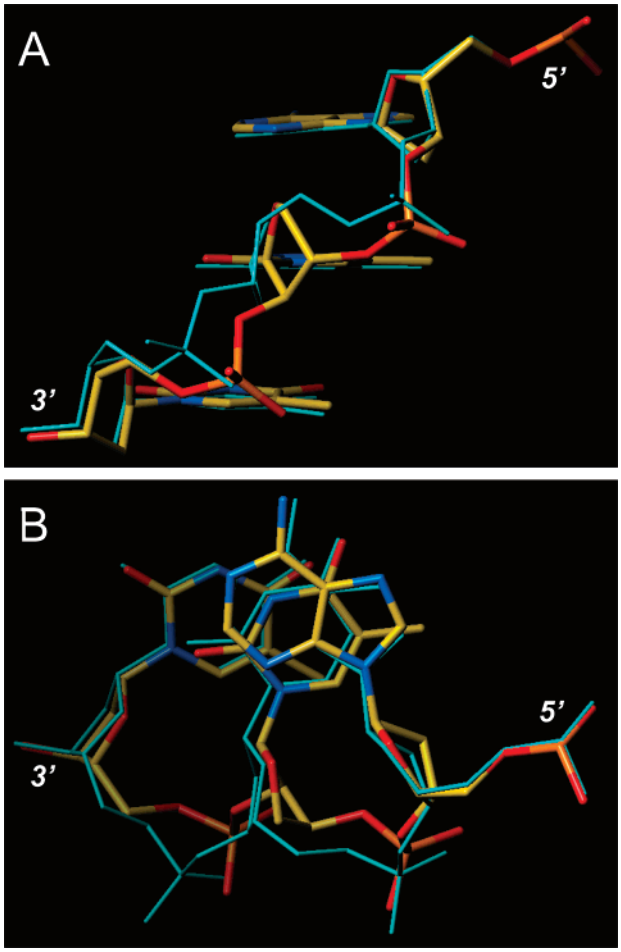


***Figure 6.*** Local alterations in the backbone geometry as a consequence of TNA incorporation. Residues A6, T*7, and T8 of the TNA-modified DDD viewed (A) perpendicular to the helical axis and (B) approximately along the helical axis. The reference duplex is included as a stick model in blue. Only base atoms of residues A6 and T8 were used to superimpose TNA-modified and reference strands.

deoxyribose O4′ atoms in the native dodecamer (Figure 3), the geometries of the hydrogen bonds between them and spine water molecules in the two structures are almost identical.

The structure of the TNA-modified duplex is an excellent example of the ability of double helical DNA to tolerate geometric perturbations in the backbone while leaving intact optimal stacking interactions. Previous high-resolution crystal structures of native DNA duplexes revealed flexibility of the backbone, albeit of a more limited nature, that appeared not to affect the base−base interactions in the core.[16−18]

## Conclusions

The 1.2 Å crystal structure of the DDD duplex containing single TNA T residues per strand reveals little difference in the global structure relative to the reference structure, demonstrating that a TNA unit can be easily accommodated in B-form DNA. The base stacking of the modified duplex is practically unchanged compared to the reference DDD. However, the

(15) Tereshko, V.; Minasov, G.; Egli, M. *J. Am. Chem. Soc.* **1999**, *121*, 3590−3595.
(16) Egli, M.; Tereshko, V.; Teplova, M.; Minasov, G.; Joachimiak, A.; Sanishvili, R.; Weeks, C. M.; Miller, R.; Maier, M. A.; An, H.; Cook, P. D.; Manoharan, M. *Biopolymers* **2000**, *48*, 234−252.
(17) Kielkopf, C. L.; Ding, S.; Kuhn, P.; Rees, D. C. *J. Mol. Biol.* **2000**, *296*, 787−801.
(18) Schuerman, G. S.; Van Meervelt, L. *J. Am. Chem. Soc.* **2000**, *122*, 232−240.

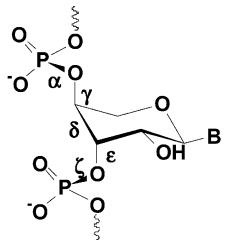

***Figure 7.*** Constitution and configuration of the repetitive unit in (L)-α-lyxopyranosyl (4′→3′) oligonucleotides.

backbones display considerable geometric changes at the sites of TNA incorporation. The tetrose sugars adopt C4′-*exo* pucker, and the resulting quasi-diaxial orientation of the 3′- and 2′-substituents allows the threose to bridge a wider gap than originally anticipated from the five backbone bonds in TNA relative to six in DNA. The minor differences in the overall duplex geometry are further corroborated by similar CD spectra obtained for the modified and reference duplexes.

Another alternative nucleic acid system that contains five covalent bonds per repeating mononucleotide unit, (L)-α-lyxopyranosyl (4′→3′) oligonucleotides (Figure 7), was also shown to pair with both DNA and RNA.[19] The model of an idealized pairing conformation of (L)-α-lyxopyranosyl nucleosides places the 4′- and 3′-phosphate groups in a diaxial orientation, similar to what is observed in the case of TNA here. This feature appears to allow both to fit almost seamlessly into B-form DNA despite the relatively large spacing between adjacent phosphate groups in B-DNA.

The present study provides detailed insight into the conformational features of TNA in a double-helical environment and allows a structural rationalization of the efficient pairing between TNA and DNA. However, it does not shed light on the geometry of TNA in an A-form duplex environment and the structural basis for pairing between TNA and RNA. We have collected preliminary crystallographic data to 1.5 Å resolution for an A-form decamer duplex containing single TNA units. Experiments to improve the resolution to a limit that is closer to that of the structure of the B-form duplex presented here are under way. Together, these crystal structures will allow a complete understanding of TNA self-pairing as well as its stable cross-pairing with both DNA and RNA.

## Experimental Section

**Materials.** The synthesis of 3′-*O*-dimethoxytrityl-(L)-α-threofuranosylthymine-2′-*O*-(2-cyanoethyl)-*N*,*N*′-diisopropyl phosphoramidite monomer was previously described.[2] 2′-Deoxynucleotide phosphoramidites and 3′-terminal nucleoside controlled pore glass (CPG) support were purchased from Glen Research (Sterling, VA). All other chemicals for solid-phase oligonucleotide synthesis were purchased from Glen Research (Sterling, VA) as well.

**Oligonucleotide Synthesis.** The native and TNA-modified oligonucleotides were synthesized on an Applied Biosystems 381A DNA synthesizer following slight modifications to published procedures.[20] Monomer coupling times were 90 s for 3′-CE phosphoramidites and 10 min for the TNA amidite. Deprotection and cleavage of the oligonucleotides from the solid support were achieved using 28% $NH_4$-OH, 55 °C for 8 h. Reverse-phase (RP) high-performance liquid chromatography (HPLC) analyses and purifications were carried out on an Applied Biosystems chromatograph with a Hewlett-Packard Hypersil ODS-5 column (4.6 × 200) and a 1% gradient of acetonitrile in 0.03 M triethylammonium acetate buffer (pH 7.0) with a flow rate of 1.0 mL/min. The crude oligonucleotides were isolated as trityl-on derivatives by RP HPLC and detritylated in 80% acetic acid for 30 min. The oligonucleotides were purified by RP HPLC a second time.

**CD Spectroscopy.** CD spectra were recorded on a JASCO J710 spectropolarimeter at 5 °C as previously described.[21] Each spectrum was the average of five scans at a rate of 0.5 nm/min. The buffer used was 90 mM NaCl, 10 mM $Na_2HPO_4$ (pH 7.0), and 1 mM EDTA with a total strand concentration of 2.5 μM. The data were processed on a PC computer with Windows based software supplied by the manufacturer (JASCO, Inc.). The CD spectra were normalized by subtraction with the reference buffer.

**Crystallization and Data Collection.** Crystals were grown at room temperature by the sitting drop vapor diffusion method. Droplets (20 μL) containing 1 mM oligonucleotide, 20 mM sodium cacodylate pH 7, 10 mM magnesium acetate, and 3 mM spermine tetrahydrochloride were equilibrated against a reservoir of 25 mL of 40% 2-methyl-2,4-pentanediol (MPD). Large hexagonal rods appeared after a week. For data collection, a crystal (0.3 × 0.2 × 0.2 mm) was picked up from a droplet with a nylon loop and transferred into a cold $N_2$ stream (120 K). High- and low-resolution data sets were collected on the 5-ID beam line ($\lambda$ = 0.978 Å) of the DuPont-Northwestern-Dow Collaborative Access Team at the Advanced Photon Source, Argonne, IL, using a MARCCD detector. Data were integrated and merged with DENZO/SCALEPACK.[22] Selected crystal data and reflection statistics are listed in Table 1.

**Structure Determination and Refinement.** The structure was solved by molecular replacement using a DNA dodecamer[11] as the search model and refined with the programs CNS[12] and SHELXL-97.[13] After the *R*-free was monitored using 10% of the reflections,[23] all reflections were included in the final rounds of isotropic refinement. Hydrogen atoms were added in SHELXL-97, all nucleic acid atoms, and many water molecules were treated anisotropically. Final refinement parameters and root-mean-square deviations from standard values are listed in Table 1. Final coordinates and structure factors have been deposited in the Protein Data Bank ID code 1N10.

**Acknowledgment.** This research is supported by the National Institutes of Health (Grant: GM 55237 to M.E.). C.J.W. is the recipient of a postdoctoral fellowship from the National Science and Engineering Research Council of Canada (NSERC). We thank Dr. K. U. Schöning and Dr. X. Wu for phosphoramidite synthesis. Portions of this work were performed at the DuPont-Northwestern-Dow Collaborative Access Team (DND-CAT) Synchrotron Research Center located at Sector 5 of the Advanced Photon Source. DND-CAT is supported by the E.I. DuPont de Nemours & Co., The Dow Chemical Company, the U.S. National Science Foundation through Grant DMR-9304725, and the State of Illinois through the Department of Commerce and the Board of Higher Education Grant IBHE HECA NWU 96. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Energy Research under Contract No. W-31-102-Eng-38.

JA0207807

(19) Reck, F.; Wippo, H.; Kudick, R.; Bolli, M.; Ceulemans, G.; Krishnamurthy, R.; Eschenmoser, A. *Org. Lett.* **1999**, *1*, 1531−1534.
(20) Wilds, C. J.; Damha, M. J. *Nucleic Acids Res.* **2000**, *28*, 3625−3635.
(21) Wilds, C. J.; Damha, M. J. *Bioconjugate Chem.* **1999**, *10*, 299−305.
(22) Otwinowski, Z.; Minor, W. *Methods Enzymol.* **1997**, *276*, 307−326.
(23) Brünger, A. T. *Nature* **1992**, *355*, 472−475.