

Solution Structure and Conformational Dynamics of Deoxyxylonucleic Acid (dXNA): An Orthogonal Nucleic Acid Candidate

Mohitosh Maiti^{[a], [f]}, Vanessa Siegmund^{[b], [f]}, Mikhail Abramov^[a], Eveline Lescrinier^[a], Helmut Rosemeyer^[c], Mathy Froeyen^[a], Amutha Ramaswamy^[d], Arnout Ceulemans^[e], Andreas Marx^[b], Piet Herdewijn^{*[a]}

Abstract: Orthogonal nucleic acids are chemically altered nucleic acid polymers not able to transfer information with natural nucleic acids and which can be used in synthetic biology to store and transfer genetic information independently. Recently it was proposed that xylose-DNA (dXNA) can be considered as a potential candidate for an orthogonal system. Here we present the solution structure and conformational behavior of two self-complementary, fully modified dXNA oligonucleotides as

determined by CD and NMR spectroscopy. These studies are the initial experimental proof for structural orthogonality of dXNA and also try to answer the question ‘why deoxyribose and not deoxyxylose’ as natural information system. In aqueous solution dXNA duplex predominantly forms a linear ladder-like (type-1) structure. This is the first example of a furanose nucleic acid that adopts a ladder-like structure. In presence of salt, type-1 duplex form exists in equilibrium with type-2 duplex form.

The corresponding nucleoside triphosphates (dXNTP) were synthesized and evaluated for their ability to be incorporated into a growing DNA chain using several natural and mutant DNA polymerases. Despite the structural orthogonality of dXNA, DNA polymerase β mutant is able to incorporate the dXNTP's showing initial DNA dependent dXNA polymerase activity.

Keywords: Xylose-DNA • Synthesis • NMR Spectroscopy • Helical Structure • Orthogonal nucleic acid

[a] M. Maiti, Dr. M. Abramov, Prof. Dr. E. Lescrinier, Dr. M. Froeyen, Prof. Dr. P. Herdewijn
Laboratory for Medicinal Chemistry, Rega Institute for Medical Research
Katholieke Universiteit Leuven
Minderbroedersstraat 10, 3000 Leuven, Belgium
Fax: (+) 32 16337340
E-mail: Piet.Herdewijn@rega.kuleuven.be

[b] V. Siegmund, Prof. Dr. A. Marx
Department of Chemistry and Konstanz Research School of Chemical
Biology
University of Konstanz
Universitätsstrasse 10, 78457 Konstanz, Germany

[c] Prof. Dr. H. Rosemeyer
Organic Materials Chemistry and Bioorganic Chemistry, Institute of
Chemistry
University of Osnabrück
Barbarastrasse 7, D-49069 Osnabrück, Germany

[d] Asst. Prof. Dr. A. Ramaswamy
Centre for Bioinformatics
Pondicherry University
Puducherry 605014, India

[e] Prof. Dr. A. Ceulemans
INPAC Institute for Nanoscale Physics and Chemistry and Quantum
Chemistry Group
Katholieke Universiteit Leuven
Celestijnenlaan 200F, 3001 Leuven, Belgium

[f] These authors contributed equally to this work

Introduction

All forms of life in nature use nucleic acids containing five-membered sugars, D-ribose (RNA) or deoxy-D-ribose (DNA) as their genetic information system. Search for alternative information systems^[1-4] might answer the questions related to, why nature has selected these systems and not others, and likewise might be important for the selection of alternative nucleic acids to build an orthogonal information system for the safe transfer of genetic information in nearly engineered life forms.^[5] This in the long run may facilitate in the creation of new microorganism for applications in medicine, energy and food supply.

Modified oligonucleotides designed in the context of antisense approach^[6] have a backbone structure mimicking the natural sugar phosphate backbone pre-organization as observed in double stranded RNA and thereby exhibit limited variability in the regular right handed duplex structures with high thermodynamic stability. In contrast, research on nucleic acid alternatives in the context of understanding chemical etiology of natural DNA/RNA system resulted in the synthesis of nucleic acids like TNA^[3, 4, 7], PNA^[8] which efficiently cross-pair with natural DNA/RNA, and self-pairing systems like homo-DNA^[1, 4, 9] and p-RNA^[1, 4, 10] having a six-membered sugar in the backbone with very high duplex stability.

A useful orthogonal nucleic acid should have a stable^[1] duplex structure with reasonable chemical and enzymatic stability, lack of cross-pairing with DNA/RNA, and preferably the backbone should have some flexibility^[4, 5] to allow formation of various secondary and tertiary structures which are needed for information storage, transfer and catalysis.

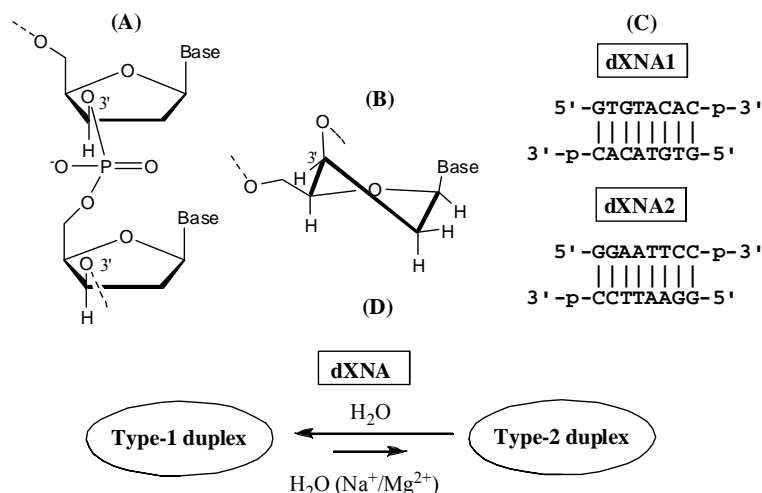


Figure 1: A) Chemical structure and view of dinucleotide step in deoxyxylose oligonucleotide (dXNA), 3'-position is configurationally altered in dXNA. B) North (N)-type puckering of deoxyxylose sugar ring as determined by NMR spectroscopy. C) Sequence and Watson-Crick base pairing in the duplex form of two self-complementary deoxyxylose oligonucleotides (dXNA1 and dXNA2) used in this study. D) Equilibrium between structurally two-different (type-1 and type-2) duplex form. In H₂O/D₂O (1.6 mM) predominantly type-1 duplex form was observed.

Xylose-DNA (dXNA) is a 3'-diastereomeric epimer of DNA (**Figure 1A**). From the structural point of view, this minimal chemical modification in the five-membered deoxyribose sugar ring leads to a change in the sugar puckering from South to North type form^[11] (**Figure 1B**), essentially changing the endo-cyclic backbone torsion angle δ and expected to implement restriction across backbone torsion γ possibly by preventing this torsion angle to adopt the usual *gauche+* conformation because of 3' O-P to 5' O-P repulsion. In view of the potential use of dXNA as an orthogonal information system, it is important to analyze how the changes of these structural parameters contribute to an alteration of the overall three dimensional duplex structure of dXNA and influence its conformational flexibility.

Recently Ramaswamy *et al.* reported the left handed helical structure and dynamics of dXNA by molecular modelling study and proposed xylose-DNA as an orthogonal nucleic acid candidate^[12]. Research on dXNA and locked-dXNA^[13, 14] was performed by Seela *et al.* and Wengel *et al.* mainly in the context of the anti-gene and antisense strategy. They synthesized deoxyxylo-derived nucleotide building blocks and their corresponding oligonucleotides and studied their nuclease stability, conformational and hybridization properties with itself and with complementary DNA^[11, 15-19] and RNA^[13, 14]. It was observed that dXNAs are stable against degradation by cellular nucleases. Notably it was observed that introduction of few deoxyxylo residues in one strand of DNA^[11, 15, 17] or RNA^[13, 14] significantly reduced thermal stability of double stranded hetero-duplexes, while complete modification of one strand abolished duplex stability^[11]. However, it was found that self-complementary fully modified dXNA can form duplexes^[11] with at least similar stability to DNA duplexes. Thermal denaturation, calorimetric and circular dichroism spectroscopic studies identified

unusual thermodynamic and conformational properties of these dXNA duplexes.^[18, 19] All these data suggest that double stranded deoxyxylose nucleic acid can adopt a left-handed duplex where the sugars adopt the thermodynamically more stable N-type sugar puckering and a right-handed duplex where the sugars adopt the less stable S-type sugar puckering. However, there are no experimental structural data available to confirm or contradict the predictions made by molecular modeling and CD spectroscopy.

Herein, we report structural studies of two self-complementary dXNA oligonucleotides (**Figure 1C**) with an alternating Purine-Pyrimidine sequence (dXNA1: 5'-GTGTACAC-3') and with a non alternating sequence (dXNA2: 5'-GGAATTCC-3'). Both dXNAs predominantly adopt type-1 duplex form in aqueous medium. With the addition of salt they exist in equilibrium between type-1 and type-2 duplex forms (**Figure 1D**). Structural and helical parameters were extracted from type-1 duplex structure and compared with the corresponding parameters of natural duplexes. These data show significant structural differences and indicate its potential for orthogonality and may contribute to the understanding of why XNA was not selected as natural genetic information system. The data are interpreted in function of the understanding of the origin of duplex formation and helicalization of dXNA duplexes and discussed in the context of self/cross-pairing tendency of dXNA. In addition, the corresponding deoxyxylose triphosphates (dXATP and dXTTP) were chemically synthesized and evaluated for their ability to be incorporated into a growing DNA chain in a template based polymerase assay. This selection system is needed in order to identify a polymerase that can be used to evolve into a DNA dependent dXNA polymerase.

Results and Discussion

Thermal (UV) melting study: Initially both oligonucleotides were evaluated for their ability to adopt secondary structure by UV-melting study. Typical reversible sigmoidal melting curves (**Figure S1 and S2**) were obtained which indicate a cooperative melting of multi stranded structural (duplex) form. Melting temperatures (T_m) for dXNA1 and dXNA2 are 56.8 and 52.2 °C (8 μ M dXNA in aqueous buffer solution) respectively at 260 nm wavelength. These melting temperatures are significantly higher than the corresponding DNA duplexes with T_m 25.8 and 27.7 °C respectively. The melting curves, however, is rather broad suggesting the existence of several structural forms in solution which melt by a non all-in-one transition.

Circular dichroism (CD) study: CD spectroscopic studies were performed for both deoxyxylo oligonucleotides in neutral (pH 7.0) aqueous solution and in neutral cacodylate-buffer solution at various temperatures in order to probe the helical form adopted by the dXNA and its dependence on solution conditions (Similar conditions were also used for NMR study). The CD spectra were highly dependent on ionic condition and temperature of the experiment, suggesting dXNA as a flexible system. dXNA1 and dXNA2 with different sequence context also behaved differently in solution, and none of their CD spectra resemble to standard CD spectrum of B form DNA^[20] (expected strong negative and positive cotton effects around 255 and 293 nm respectively) and A form RNA^[21] (expected strong negative and positive cotton-effects around 212 and 258 nm respectively).

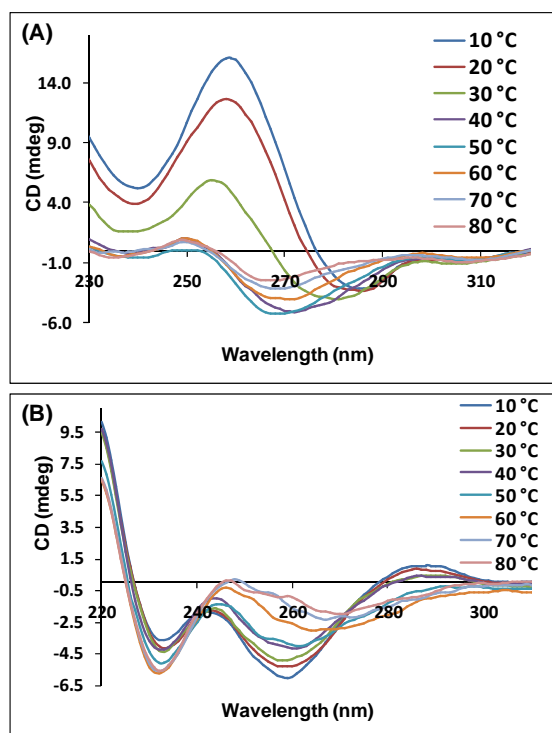


Figure 2: Temperature dependent CD spectra of dXNA1 oligonucleotide, (A) in H₂O (12 μ M) at pH 7.0 and (B) in aqueous buffer solution [containing sodium cacodylate (10 mM), NaCl (100 mM), MgCl₂ (10 mM)] of pH 7.0.

CD spectrum of dXNA1 in water at 10 °C (**Figure 2A**) showed an intense positive Cotton effect at 258 nm with a small negative band at 285 nm. This spectrum resembles the mirror-image^[11] of right-handed B-DNA form and close to the natural Z-DNA spectrum^[20], indicating the existence of a left-handed structural form. With the increase of temperature from 10 to 40 °C the positive Cotton effect gradually disappears with a concomitant hypsochromic shift (to 249 nm) along with a hypsochromic shift (to

271 nm) of the small negative band, this may indicate structural change to a predominantly non-helical (or slightly right-handed) form. In contrast, CD spectrum of dXNA2 in water at 10 °C (**Figure S3**) showed a moderately intense negative Cotton effect at 263 nm with a small positive band at 240 nm, which may be indicative for a slightly right-handed (or non-helical) form. With the rise of temperature up to 30 °C the negative Cotton effect increases, indicating a change in helicity, and then up to 80 °C gradually decreases showing a typical melting effect.

CD spectrum of dXNA1 in buffer solution at 10 °C (**Figure 2B**) is unusual showing two negative bands at 231, 258 nm and a less intense positive band at 287 nm. Similar CD spectra were previously reported for dXNA with an alternating sequence^[19]. With increase in temperature (up to 80 °C) the bands at 258 and 287 nm merged together and appeared as a broad less intense negative band around 269 nm. CD spectrum of dXNA2 in buffer solution at 10 °C (**Figure S4**) showed two negative bands at 231 and 265 nm. With temperature rise (up to 80 °C) the 265 nm negative band decreases in its intensity and a new small positive band appears at 242 nm. These complex CD spectra in presence of buffer cannot be directly related to one type of known helical structural form. This may result from the existence of equilibrium between several helical and/or non helical structural forms.

NMR spectroscopic study

(A) Identification of two types of duplex form: Signals for the major structural form (type-1) of both dXNA in aqueous medium were observed in the NMR spectra (**Figure S5**). Initially, the existence of a second minor structural form was detected by the appearance of very small peaks close to the base line of proton spectra in D₂O and by the observation of weak exchange cross-peaks in 2D NOESY spectra in D₂O or in H₂O-D₂O (9:1) (**Figure S6**). Titration of the samples with solutions of NaCl and MgCl₂ in D₂O gradually triggered the formation of second structural form as observed in 1D proton (**Figure S5**) and phosphorus spectra (**Figure S7**).

Formation of hydrogen bonded duplex form (type-1) was identified by the appearance of three sharp imino-signals in the proton spectra recorded in H₂O-D₂O (**Figure S8A, S9A**). In principle, as the sequences are self-complementary, four hydrogen bonded imino-protons should be observable in an 8-mer Watson-Crick anti-parallel duplex. The terminal imino proton was not observed, as expected by the typical end-fraying effect. 1D imino-proton spectra (**Figure S8B, S9B**) for both dXNAs with high salt (Na⁺/Mg²⁺) concentration showed highly shielded 3 additional strong imino-signals (for type-2 duplex) apart from the imino-signals observed in water (type-1), suggesting equilibrium between type-1 and type-2 duplex form as shown in **Figure 1D**.

Diffusion measurement by DOSY-NMR in D₂O with high salt concentration at 10 °C showed (**Figure S10**) faster diffusion of type-2 duplex ($\sim 8.83 \times 10^{-11}$ m²/s) than type-1 duplex form ($\sim 6.46 \times 10^{-11}$ m²/s), indicating more compact structure of type-2 compared to type-1 duplex.

(B) Watson-Crick base pairing in duplex forms: Hydrogen-bonded imino proton peaks for type-1 dXNA1 and dXNA2 duplex are well within expected chemical shift range of standard Watson-Crick base pairing in DNA/RNA duplex^[22]. Imino-protons of thymine residues are deshielded from the guanine residues as expected. For A-T base pairing steps strong NOE interaction from H2 proton of adenine to the hydrogen-bonded imino proton of thymine (also assisted to identify A-T pairing and in the assignments of H3-T protons) and NOE interaction for the imino-proton of thymine to the amino protons of adenine were observed (**Figure S11, 12**). In G-C steps expected NOE interactions between the imino proton of guanine residues and amino protons of cytosine residues

and weak spin diffusion cross-peak to cytosine H5 proton were observed (Figure S11, 12). Classical imino to imino proton walk was visible (very weak) in Watergate NOESY spectrum of dXNA1 and 2. Moreover, as expected^[22] the chemical shifts of C4 carbons of thymine bases (~ 167.7 ppm) were deshielded due to hydrogen bonding of C4-carbonyl oxygen in the A-T base pairing step, while C2 carbons (~ 150.5 ppm) of the free carbonyl groups of thymine bases were in up-field (Figure S13).

In type-2 duplex generally all imino-protons are shielded when compared to the imino-protons of the type-1 duplex. All imino protons of type-2 duplex are in conformational exchange (Figure S6) with type-1 duplex as exchange cross-peaks were observed in NOESY spectra (in H₂O-D₂O with salt).

(C) Resonance assignments of type-1 duplex: All non-exchangeable and exchangeable protons of type-1 duplex structure of dXNA1 and dXNA2 were assigned completely.

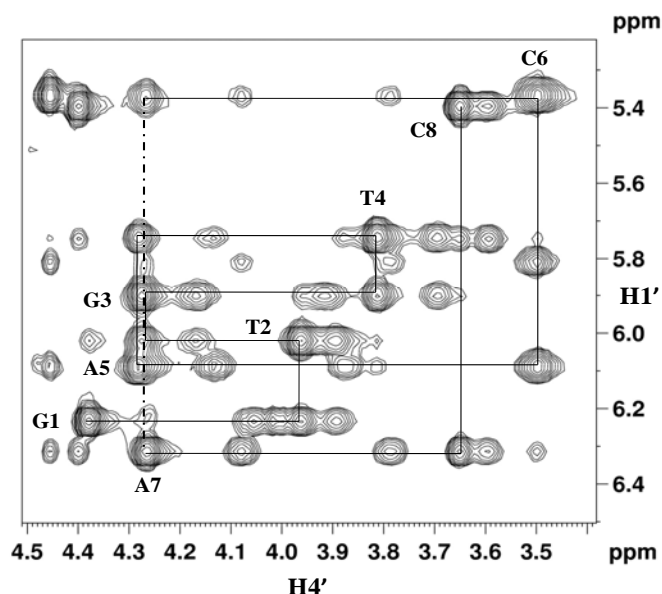


Figure 3: Expansion of NOESY spectrum in D₂O (150 ms, at 10 °C) for dXNA1 (5'-G'TGTACAC-3') duplex, showing a clear NOE walk between H1'-H4' intra-residue and H1'(n)-H4'(n+1) inter-residue protons.

(1) Non-exchangeable protons: For dXNA1, the standard NOE-walk^[23] as in DNA between aromatic (H8/H6) protons of base and anomeric (H1') sugar protons was absent, since only intra residue H8/H6-H1' NOE cross-peaks were observed but no inter residue H1'(n)-H8/H6 (n+1) NOEs were visible. However, H4' protons of dXNA1 were sufficiently resolved and a clear NOE-walk (Figure 3) between intra residue H1'-H4' NOEs and inter-residue H1'(n)-H4'(n+1) NOEs was observed providing sequential assignments of successive sugar residues from NOESY (150 ms) spectrum. H2', H2'' and H3', H8, H5, H6, H7 (Me-T) protons of individual residues were assigned via correlations observed in DQF-COSY (Figure S14), TOCSY and NOESY spectrum. The H2 protons of adenine bases were assigned by strong NOE correlations between the H1' proton of the next residue (n+1) from the same strand and the H1' proton of the base-paired partner of the previous residue (n-1) from the opposite strand (same strand in self-duplex). H5' and H5'' protons are assigned based on the observed very strong intra residue NOE correlation with the aromatic H8/H6 protons, suggesting *anti* orientation of nucleobases and *gauche* to *antiperiplanar* orientation of the γ torsions. In this orientation of γ torsion angle, the H5'' protons of each residue will be situated close to the previous residue nucleobase. This generally will result in the shielding of H5'' protons with respect to the H5' protons which was indeed observed.

Notably in the pyrimidine steps the H5'' protons were highly shielded, providing stereospecific assignments of H5' and H5'' protons.

Additionally, natural abundance ¹H-¹³C HSQC spectrum for sugar and base region facilitated the assignment process by validating and clarifying overlapping ¹H signals. All phosphorus signals were assigned by ¹H-³¹P correlations in HETCOR spectrum (Figure S15, 16), which finally reconfirmed the assignments of proton signals of successive backbone residues.

In dXNA2, H4' proton-region are highly crowded by signal overlaps, hence, analogous procedures as used for proton assignments in dXNA1 could not be followed. However, a weak less pronounced similar (with additional inter-residue H6/H8 (n)-H1'(n+1) NOEs) to standard NOE-walk (Figure S17) was observed in NOESY spectrum with higher mixing time (300 ms), providing connectivity between residues. For assignments of all other protons an analogous treatment of dXNA2 spectra as in dXNA1 was followed.

(2) Exchangeable protons: All exchangeable protons of type-1 duplex of dXNA1 and dXNA2 were assigned as described in the previous sections.

(D) Structural parameters from spectra of type-1 duplex:

(1) Deoxyribose sugar puckering: All five-membered deoxyribose sugar rings adopt North (N)-type (3'-endo) sugar puckering (Figure 1B) as determined from the vicinal ³J_{H1'-H2'} and ³J_{H1'-H2''} proton-proton couplings. ³J_{H1'-H2'} coupling was around 0 Hz for all residues, as indicated by the appearance of H1' proton peaks as doublets (³J_{H1'-H2''} = 4.5 to 7 Hz) and by the absence of ³J_{H1'-H2'} cross-peaks in DQF-COSY spectra (Figure S18). However, in NOESY and in TOCSY spectra H1' protons gave cross-peaks with both H2' and H2'' protons, this facilitated in the stereospecific assignment^[23] of H2' and H2'' protons of dXNA1 and dXNA2.

(2) Backbone torsions: The β torsion angles (C4'-C5'-O5'-P5') were restrained to *trans* region in structure calculations as determined by experimental vicinal hetero-nuclear phosphorus proton (³J_{P5'-H5'}, ³J_{P5'-H5''}) couplings (Table S1) derived from ³¹P coupled and decoupled DQF-COSY spectrum (and based on standard Karplus plot^[24]). For purine and pyrimidine nucleotides in the dXNA1 duplex the β torsion angle range 180±75° and 200±40° were restrained respectively. In the dXNA2 duplex, for all nucleotides β range 180±40° was applied.

The γ torsion angles (C3'-C4'-C5'-O5') were restrained to *gauche* to *antiperiplanar* region in structure calculations as determined by vicinal proton-proton (³J_{H4'-H5'}, ³J_{H4'-H5''}) couplings (Table S1) derived from ³¹P decoupled DQF-COSY spectrum (Figure S14)^[24]. For purine and pyrimidine nucleotides of dXNA1 duplex the γ torsion angle region 210±35° and 270±45° were restrained respectively. In dXNA2 duplex, for all nucleotides γ 210±35° were restrained except for T5 residue in the junction between purine and pyrimidine sequences (γ = 250±30°).

The ϵ torsion angles (P3'-O3'-C3'-C4') were restrained to *trans* region (ϵ = 180±60°) in all residues of dXNA1 and dXNA2 duplex based on the observed cross-peaks for long range phosphorus proton couplings (⁴J_{H4'-P3'}) in ¹H-³¹P HETCOR spectra (Figure S15, 16), indicative for W-shaped conformation of P3'-O3'-C3'-C4'-H4' fragment in the backbone.

Solution structure of type-1 duplexes: The structures of modified duplexes (type-1) were determined by performing torsion angle molecular dynamics followed by the refinement of 20 selected structures using NMR-derived distance and dihedral restraints in X-PLOR NIH program^[25] (see experimental section). During simulation, the structures converged to a family of structures with similar geometry and energy. The experimentally determined restraints were implemented to retain inter-proton distances,

backbone torsion angles, and sugar pucker of all residues. Structure determination statistics are provided in **Table S2**.

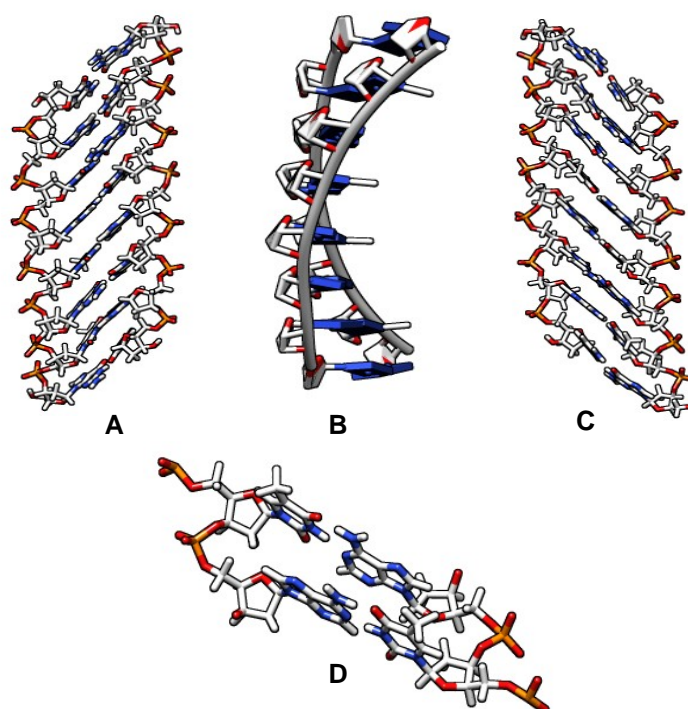


Figure 4: Solution structure of dXNA1 self-complementary type-1 duplex, (A) Minor groove view, (B) Central view with ribbon representation, (C) view at major groove face, and (D) a dinucleotide step showing inter-strand stacking interaction.

Table 1: Torsion angles and diagnostic helical parameters for natural B-DNA and A-RNA duplexes, left-handed (L-dXNA/xb29_md40) and intermediate (I-dXNA/xa13_md20) structural model of dXNA duplexes derived from MD-simulation, and the NMR solution structures of two type-1 dXNA duplexes under this study. Standard values for DNA, RNA are from ref.^[45] and values for L-dXNA, I-dXNA from ref.^[12] Average and standard deviations (\pm) were calculated from out-put data for an ensemble of 20 refined NMR structures using CURVES program. Backbone-base inclination angles (η_b) were calculated using 'inclination' program (ref.^[30]). For all helical parameters, see Table S3 in the supporting information.

Torsion angles ($^\circ$)	B-DNA	A-RNA	L-dXNA	I-dXNA	dXNA1	dXNA2
α	-47	-62	87	75	53 \pm 18	68 \pm 13
β	-146	180	-148	-163	-156 \pm 14	-157 \pm 10
γ	36	47	-160	-116	-122 \pm 8 (A/G)	-108 \pm 13
					-64 \pm 4 (C/T)	-77 \pm 1 (T5)
δ	156	83	-36	-28	-25 \pm 3	-24 \pm 2
ϵ	155	-152	126	147	146 \pm 13	144 \pm 10
ζ	-95	-73	71	72	75 \pm 12	67 \pm 8
χ	-98	-166	-85	-157	-161 \pm 6	-155 \pm 5
Helical parameters						
Twist ($^\circ$)	36.0	32.7	-20.0	-6.6	2.7 \pm 0.5	3.5 \pm 0.6
Inclination ($^\circ$)	-5.9	15.8	24.0	-49.0	-52.2 \pm 0.4	-49.0 \pm 0.5
Inclination- η_b ($^\circ$)	0	-30	-31	38	51	49
Rise (\AA)	3.4	2.8	4.0	5.0	5.8 \pm 0.06	5.1 \pm 0.05

Inspection of the structures for dXNA1 (**Figure 4** and **S19**) and dXNA2 (**Figure S19**) shows that both duplexes adopt a very similar overall geometry. The duplexes adopt slightly right handed, albeit quasi-linear extended ladder-like structure with significantly impaired helicity. As both duplexes are rather short (8 mer), it was not possible to quantify helical pitch and residues per turn, but they are increased in comparison with regular DNA duplexes, and this widens both major and minor grooves making them almost indistinguishable. In all duplexes, the deoxyxylose sugar residues adopt N-type sugar pucker mode. Base pairs in the duplex are

highly exposed to solvent, and strongly inclined with respect to sugar phosphate backbone. A comparison of torsion angles and helical parameters of type-1 duplex of dXNA1 and dXNA2 with that of the left-handed (L-dXNA) and intermediate (I-dXNA) xylose-DNA duplex calculated in a previously reported molecular dynamics study^[12] and standard B-DNA and A-RNA helices are given in **Table 1**. It is clear from the values of Table 1, that the parameters are close to I-dXNA duplex form and to some extent close to L-dXNA, but significantly different from those of standard A- and B-type helices. Variation of torsion angles in individual residues of the

final set of 20 calculated structures of dXNA1 and dXNA2 are provided in the supplementary **Figure S20** and **S21** respectively.

In order to validate the determined structural model of type-1 duplex derived from NMR-restraints based on NOE and coupling data, we calculated the chemical shifts for the non-exchangeable protons (**Table S4**) from the minimum energy dXNA1 duplex structure using the Nuchemics program^[26] and compared them with the experimental chemical shifts. Generally, the majority of the calculated shifts are close to experimental values, and notably for H5' and H5'' protons of purine residues a high agreement in values was observed. For pyrimidine steps, the high shielding effect of H5'' protons are confirmed by the calculated values, as also experimentally detected up to the 1.6 ppm shielding from H5' peak.

Comparison of structural parameters of dXNA with other duplexes: Torsion angles and helical parameters of dXNA1 and dXNA2 are largely identical, with only exception for the torsion angle γ , which adopts different angles for purine (-122 °) and pyrimidine (-64 °) residue of dXNA1, and T5 residue (-77 °) in the only purine-pyrimidine step of dXNA2.

Parameters of dXNA1 and dXNA2 in comparison with L-dXNA model are similar except for a slightly negative twist (-6.6 °) in the latter. On the other hand left-handed helical L-dXNA model duplex differs from dXNA1 and dXNA2 by large negative twist (-20 °, typical parameter defining left-handedness), inclination (24 °), rise (4 Å), and mainly by torsion angles γ and χ .

Comparing the backbone torsions of dXNA1 and dXNA2 with respect to natural DNA, RNA backbone show almost inverted mirror-image relationship (opposite in sign and magnitude) in α (*gauche+* vs. *gauche-*), γ (*gauche-* to *anticlinal-* vs. *gauche+*), δ (*synclinal-* vs. *gauche+* to *anticlinal+*) and ζ (*gauche+* vs. *gauche-*) except for β , ϵ , and χ , which are in *anti* domain in both natural and modified duplexes. This demonstrates chiral orthogonality of dXNA backbone in comparison with natural DNA, RNA backbone.

Comparison of helical parameters of dXNA1 and dXNA2 with natural duplex shows remarkable difference in helical twist (2-3 vs. 32-36 °), base-pair rise (5-6 vs. 2.8-3.4 Å) and inclination (-49 to -52 ° vs. 16 to -6 °). A large positive twist and lower base-pair rise in natural duplexes are the typical hallmark of compact right-handed duplexes with predominant intra-strand base stacking interaction^[27], in contrast to the extended dXNA duplexes showing large base-pair inclination. In the calculated dXNA duplexes, this large base-pair inclination induces preferential inter-strand zipper-like base stacking as inter-strand diagonal bases are close to ideal stacking distance of ~3.4 Å in type-1 duplex structure as shown in **Figure 4D**. Previously, six-membered sugar modified DNA (homodNA) and RNA (p-RNA) analogues^[4, 10] were found to adopt this kind of quasi-linear duplex structure with high inclination and inter-strand base stacking interaction in solution. Acyclic minimal nucleic acid (Glycol nucleic acid^[28], GNA) also adopts extended helical duplex with high inclination and inter-strand base stacking^[29]. The predominant inter-strand stacking interaction may contribute to the higher thermodynamic stability of the dXNA duplexes. The higher thermodynamic stability of double stranded dXNA in comparison with dsDNA is a disadvantage, since it will be more difficult to unwind to make the genetic information available.

Generally, structural parameters like intra-strand inter-phosphate distances^[28, 29] and backbone-base inclinations^[28, 30] (η_b) are used as predictive tool to rationalize potential cross-pairing behavior of modified nucleic acids with other natural or modified nucleic acids. Intra-strand P-P distance in dXNA1 (5.5± 0.7 Å) and dXNA2 (5.7± 0.2 Å) duplex are close to P-P distance in RNA (5.9 Å) than DNA (7.0 Å) duplex. However, backbone-base inclination of dXNA1 (51 °) and dXNA2 (49 °) are significantly different from DNA (*ca.* 0 °) and RNA (*ca.* -30 °), which may explain the difficulty of dXNA strands for cross-pairing with natural DNA/RNA strands, as previously reported by thermal melting studies^[11, 14].

Molecular Dynamics Simulation: MD simulations up to 20 nsec have been performed on both dXNA duplexes, using different starting conformations and simulation temperatures (283, 300 K: **Table S5**). All simulations are in explicit solvent. Resulting geometries were also compared (**Table S6-8**) to previous calculations on dXNA with 13 and 29 base pairs^[12]. The data clearly point to the existence of two alternative structures: an almost linear ladder-like structure as shown in **Figure 4** and a left-handed helical structure (**Figure S22**).

Starting from the NMR structure, dXNA1 remains in an NMR-like conformation with almost zero helicity (twist angle $0 \pm 2^\circ$), and a strong inclination of about -45°. This alternating duplex may also exist in a left-handed form with a typical negative helical twist around -30°, but a clearly distinct inclination range between +30° and -10° depending on the temperature. For dXNA2 a similar structural variety is found, but in this case the NMR structure is more flexible and the equilibrium between both structures is clearly shifted to the left-handed helix. In all cases the sugar puckering adopts C3'-endo conformation, in line with the NMR model. The glycosyl torsion angle, χ , amounts to -160° in the NMR structure, and the typical anti-conformation of -90° in the left-handed helix. Solvent exposure certainly has an influence on the relative stability of both conformations. In the previous study for the 13-mer, simulations in explicit solvent stabilized the ladder-like structure, while implicit solvent led to the left-handed helix.

The MD simulations confirm the stability of the duplexes and reproduce the characteristic strong inclination of the ladder structure, but they also predict the existence of an equilibrium with the left-handed form. Both the NMR results and the complex CD spectra seem to support the existence of such an equilibrium, but identification could not be carried out due to severe signal overlap from two different conformations and due to exchange broadening of the NMR signals.

Results of primer extension assays: We explored the action of xylotymidine and 2'-deoxyxyloadenosine 5'-triphosphates (dXTTP and dXATP) on DNA polymerases and investigated the propensity of different DNA polymerases to accept the deoxyxylose triphosphates in primer extension experiments. It was observed that all tested DNA polymerases of eukaryotic, prokaryotic and archaic origin accept dXTTP and dXATP as building blocks in primer extension reactions, except for KlenTaq DNA polymerase (**Figure 5**). Besides several commonly used DNA polymerases, we also tested DNA polymerase mutants with broadened substrate spectra that had been generated by either directed protein evolution or by site-directed mutagenesis (DNA polymerase β 5P20^[31], Terminator DNA polymerase M3^[32], KlenTaq DNA polymerase DM^[33], Taq DNA polymerase M1^[34]). Due to the different configuration of the 3'-hydroxyl group of the synthesized deoxyxylose triphosphates that is required for DNA strand elongation after incorporation, we investigated how many of the xylose-nucleotides can be incorporated in a row. Therefore, we used two DNA templates bearing only 2'-deoxyadenosine or thymidine residues after the primer binding site to call for the multiple incorporation of the respective xylo-nucleotide. A 23-nucleotide ³²P-labeled primer was used along with these templates. Subsequently, all reactions were analyzed by denaturing polyacrylamide gel electrophoresis (PAGE) and phosphorimaging.

We performed control experiments without any dNTP (**Figure 5**; lane 1) and in the presence of natural TTP or dATP, respectively. We found for all tested DNA polymerases that primer extension reactions where natural TTP or dATP was added resulted in the formation of the full-length product (**Figure 5**; lane 2). When we replaced the natural TTP or dATP by either dXTTP or dXATP, we could observe that all of the tested DNA polymerases were able to incorporate one deoxyxylose-nucleotide, except for KlenTaq DNA polymerase, and that most of them were also able to further extend this first incorporated xylose-moiety with another modified nucleotide.

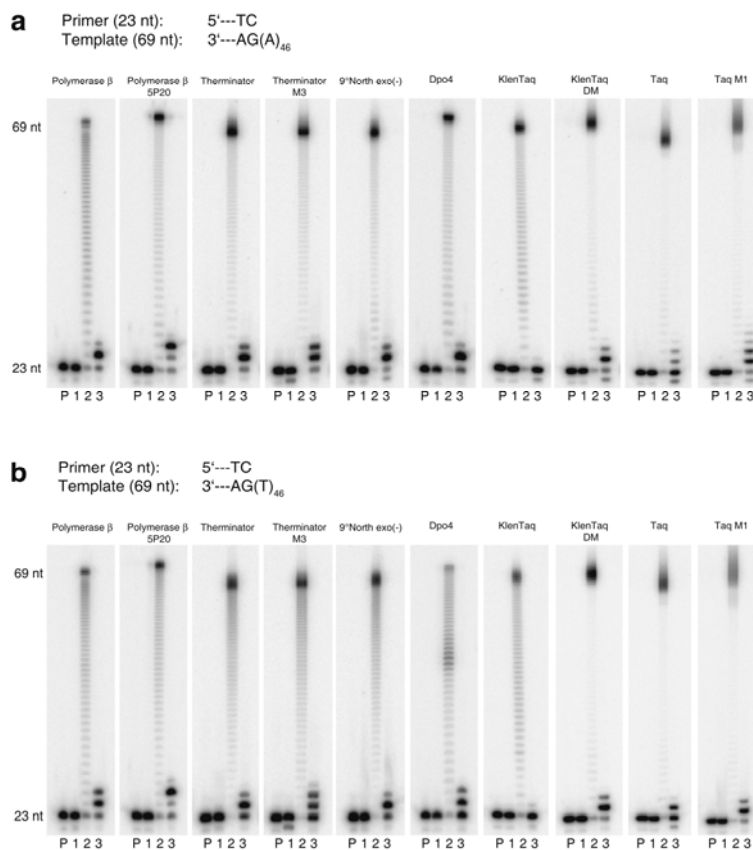


Figure 5: Primer extension studies using a) xylothyminine and b) 2'-deoxyxyloadenosine 5'-triphosphate (dXTTP and dXATP) as substrate. Partial sequences of the primer and templates are indicated on top. The employed DNA polymerases are displayed above each panel. P: ³²P-labeled primer; lane 1: control reaction without any dNTP; lane 2: primer extension performed in the presence of a) TTP or b) dATP; lane 3: primer extension performed in the presence of a) dXTTP or b) dXATP.

Additionally, we found that the tested mutants show improved extension behavior in comparison to the respective wild-type DNA polymerase. The KlenTaq DNA polymerase mutant DM was found to incorporate one xylose-nucleotide and even to extend this (**Figure 5a, b**; KlenTaq DM; lane 3) whereas the wild-type KlenTaq DNA polymerase did not show any activity in the presence of dXTTP or dXATP. The Taq DNA polymerase mutant M1 seems to be most proficient in extending the primer with up to four xylothyminine building blocks (**Figure 5a**; Taq M1; lane 3). The human DNA polymerase β mutant 5P20 was found to be most efficient in incorporating one deoxyxylose-nucleotide and to extend this for another deoxyxylose-moiety (**Figure 5**; Polymerase β 5P20; lane 3). Surprisingly, DNA polymerase Dpo4, a member of the error-prone, Y-family, TLS polymerases, was found to be just moderate efficient in extending the first incorporated xylose-nucleotide (**Figure 5a, b**; Dpo4; lane 3).

The relatively poor substrate property of deoxyxylonucleoside triphosphates compared to dNTPs against natural DNA primer:template complex and polymerases suggests orthogonality of dXNA. On the other hand, the low level of initial incorporation as shown by some mutant DNA polymerases indicates the possibility for using directed protein evolution to generate efficient DNA dependent dXNA polymerase. This can help to copy existing genetic information from a DNA template to a dXNA episome for further evaluation in synthetic biology as orthogonal information system.

Conclusion

Using NMR and CD spectroscopy, we have identified the solution state conformation of xylose-DNA. The experimental data demonstrate that the inversion of chirality of only one stereogenic center of all the five-membered residues in the sugar phosphate backbone of DNA induces significant structural alteration while increasing the flexibility of the dXNA structure.

The presence of a 2'-hydroxyl group in the ribose sugar of RNA (*vs.* deoxyribose sugar in DNA) is known to favor 3'-endo sugar puckering. This is also the preferred conformation of the deoxyxylose sugar ring in dXNA. In formose reaction, ribose is the kinetically controlled predominant product, while xylose is a thermodynamically stable minor product^[35]. This suggests that xylose nucleic acid and its deoxy congener could have also been selected by nature as information systems. There are, however, several properties of the XNA/dXNA system which makes it less suitable for these purposes. In the N-type and S-type sugar puckering, the 2'-hydroxyl group and 3'-O-phosphate group of RNA have both an axial-equatorial relationship suitable for intra molecular catalysis (for example in phosphodiester bond cleavage). This is not the case for XNA where both groups are axially oriented in the N-type conformation. The deoxyxylonucleic acid duplex is conformationally more flexible than DNA as shown by CD studies and molecular modeling and thermodynamically more stable than corresponding DNA duplex (as observed from T_m data)^[18], which makes it more difficult to be used as information system and thereby giving double stranded dXNA an entropic disadvantage in molecular

interactions. As far as the structural stability and compactness are concerned, ribose based natural DNA/RNA information system is superior to dXNA.

Our structural data of dXNA type-1 duplex showed that the conformation of the double stranded form is significantly different from that of natural nucleic acids. This may suggest that dXNA can be considered as a potential candidate for evaluating as an orthogonal nucleic acid. However, structure-based prediction for cross-pairing behavior has its inherent limitation^[30]. Since natural nucleic acids also show structural polymorphism and significant structural plasticity depending on length, sequence and environmental conditions, detailed cross-pairing studies are necessary with mixed-mer dXNA sequences targeting RNA strands in order to further investigate the orthogonality of dXNAs. So far only very limited literature data are available for cross-pairing of dXNA with RNA^[13, 14]. From structural point of view, it would be important to further investigate the inherent flexibility of dXNA backbone by designing a suitable self-complementary sequence, and to study the formation of a hairpin form.

Primer extension studies of dXTTP using different polymerases demonstrate up to two nucleotide incorporation in most of the cases and interestingly with some mutant polymerases even up to four nucleotides. Despite the structural difference of dXNA system, initial DNA dependent XNA polymerase activity has been demonstrated which opens the possibility to start with directed protein evolution in order to improve incorporation efficiency, and to subsequently create a DNA dependent dXNA polymerase.

Experimental Section

Experimental details and procedures for dXNA oligonucleotide synthesis, UV-melting experiments, CD experiments, NMR experiments, and MD simulations are provided in the supporting information.

NMR spectroscopy-derived restraints: Distance restraints were derived from NOESY spectra recorded in D₂O with 50, 100, and 150 ms mixing times. By using the computer-aided resonance assignment (CARA) program^[36], the cross-peaks in NOESY spectra were assigned and the cross-peak volumes were obtained by integration. Based on the NOE buildup curves, inter-proton distances were calculated. The calibration of NOE cross-peak intensities was performed against the H5-H6 cross peaks as an internal standard. An experimental error ($\pm 20\%$) was used on the calculated distances during structure calculation. NOE cross-peaks arising from exchangeable protons in the WATERGATE NOESY spectrum were classified as strong, medium and weak peaks; corresponding distance restraints of 1.8-3.0, 1.8-4.0 and 1.8-5.5 Å were used respectively. Hydrogen-bonding restraints were used as distance restraints in all structure calculation.

All deoxyribose sugar rings were restrained to North (N)-type sugar puckering in structure calculations [dihedral restraints applied: H1'-C1'-C2'-H2' (99.2 \pm 20°), H2'-C2'-C3'-O3' (35 \pm 25°), O3'-C3'-C4'-H4' (-157 \pm 25°)].

Backbone torsions β , γ and ϵ were restrained as mentioned in the results and discussion section under backbone torsions. No torsion angle restraints were implemented on the backbone dihedral angle α , and ζ in any structure calculations.

Duplex symmetry and base-pair planarity restraints were implemented in the initial stages of structure calculation.

Structure calculations (Type-1 duplex of dXNA1 and dXNA2): All structure calculations were performed using the X-PLOR NIH program^[25]. The standard topology and parameter file (topallhdg.dna, parallhdg.dna) for DNA were adapted for deoxyribose residue by introducing inversion in the 3'-chiral improper angle. The torsion-angle molecular dynamics protocol used was largely identical to that proposed for the DNA duplex. A set of 100 structures was generated by torsion-angle molecular dynamics, starting from two extended strands with random structures by using all structural restraints. After the torsion angle molecular dynamics round, the majority of the structures were converged to very similar structures with similar total energies and with no violations of the NOE and dihedral restraints. The 20 lowest energy structures were used for further refinement (in vacuum, and in a minimized water box) during a gentle molecular dynamics round. The final refinement was started with 20 ps constant-temperature molecular dynamics simulation at 300 K (20000 steps of 0.001 ps) and was followed by a 200-step conjugate-gradient energy minimization on an average structure of the last 10 of 20 ps molecular dynamics simulation. Torsion angles and helical parameters of the obtained solution structures were determined using CURVES 5.3^[37].

Nuclechmics program^[26] was used to back-calculate proton chemical shifts (Table S4) considering the contribution of aromatic ring-current and magnetic-anisotropy effect using the minimum energy dXNA1 duplex (type-1) structure. Finally, some visual representations of the molecules were obtained using the UCSF Chimera 1.4.1 program.

Synthesis of xylothymidine and 2'-deoxyxyloadenosine 5'-triphosphates:

Xylothymidine and 2'-deoxyxyloadenosine were synthesized starting from the corresponding natural nucleosides following the previously reported methods^[15]. Configurationally altered 3' hydroxyl group was selectively protected by benzoyl group^[38]. Finally one pot synthesis of Ludwig^[39] was applied to obtain the corresponding deoxyxylo nucleoside 5'-triphosphates (A and B). The removal of 3'-O-benzoyl group was performed with aqueous ammonia. Triphosphates were purified by anion exchange chromatography.

A) 1-(2'-deoxy- β -D-xylofuranosyl) thymine 5'-triphosphate triethylammonium salt:

To a colorless solution of 3'-benzoyl protected xylothymidine (100 mg, 0.352 mmol) in pyridine (2 mL) POCl₃ (0.049 mL, 0.528 mmol) was added. The reaction mixture was stirred for 2h under argon at 4 °C and turn to yellow color. A solution of (NBu₄)₂H₂P₂O₇ (1404 mg, 2.64 mmol) in DMF (5 mL) and Bu₃N (0.652 mL, 3.52 mmol) was added and after stirring for 2 min the reaction mixture was poured into cold 1M TEAB (5 mL). The reaction mixture was concentrated under reduced pressure and the reaction product was deprotected with 25% aqueous ammonia for 3h. Purification of the triphosphate compound was performed on a Source 15Q ion exchange column (Amersham Biosciences) using a TEAB gradient from 0 to 0.5 M (pH 7.5) in 30 minute. ³¹P NMR (D₂O): δ -6.56 (d, 1 P, α), -11.23 (d, 1 P, γ), -22.79 (t, 1 P, β). MS (ESI): Calcd: 480.9814 [M-H]. C₁₀H₁₇N₅O₁₄P₃. Found: 480.9815 [M-H]

B) 9-(2'-Deoxy- β -D-xylofuranosyl) adenine 5'-triphosphate triethylammonium salt:

To a colorless solution of 3'-benzoyl protected 2'-deoxyxyloadenosine (80 mg, 0.225 mmol) in trimethylphosphate (2 mL) POCl₃ (0.049 mL, 0.528 mmol) was added. The reaction mixture was stirred for 2h under argon at 4 °C and turn to yellow color. A solution of (NBu₄)₂H₂P₂O₇ (900 mg, 1.7 mmol) in DMF (3 mL) and Bu₃N (0.652 mL, 3.52 mmol) was added and after stirring for 2 min the reaction mixture was poured into cold 1M TEAB (5 mL). The reaction mixture was concentrated under reduced pressure and the reaction product was deprotected with 25% aqueous ammonia for 3h. Purification of the triphosphate compound was performed on a Source 15Q ion exchange column (Amersham Biosciences) using a TEAB gradient from 0 to 1 M (pH 7.5) in 45 minute. ³¹P NMR (D₂O): δ -10.88 (d, 1 P, α), -11.20 (d, 1 P, γ), -23.26 (t, 1 P, β). MS (ESI): Calcd: 489.9930 [M-H]. C₁₀H₁₆N₅O₁₂P₃. Found 489.9929 [M-H]

Natural and mutant polymerases, primers and templates: The recombinant enzymes and mutants were expressed and purified as described in literature (human DNA polymerase beta: adapted from ref.^[40]; Dpo4 DNA polymerase (*Sulfolobus solfataricus* P2 DNA polymerase IV): adapted from ref.^[41]; Terminator DNA polymerase (D141A, E143A, A485L mutant of *Thermococcus species* 9°N-7 DNA polymerase): see ref.^[32]; 9°North exo(-) DNA polymerase (3'→5' exonuclease-deficient variant of *Thermococcus species* 9°N-7 DNA polymerase): adapted from ref.^[42]; KlenTaq DNA polymerase (Klenow Fragment of *Thermus aquaticus* DNA polymerase): see ref.^[43]; Taq DNA polymerase (*Thermus aquaticus* DNA polymerase): see ref.^[44]). The purity of the proteins was > 95% as controlled by SDS-PAGE. The enzyme concentrations were determined by Bradford assay. Unmodified TTP and dATP were purchased from Roche. DNA oligonucleotides were from Purimex, Germany, and double HPLC-purified.

Primer extension assays: The reaction mixtures (20 μ L) contained 150 nM ³²P-labeled primer (5'-GAC CCA CTC CAT CGA GAT TTC TC-3'), 225 nM of either polyA-template (3'-CTG GGT GAG GTA GCT CTA AAG AG(A)₄₆-5') or polyT-template (3'-CTG GGT GAG GTA GCT CTA AAG AG(T)₄₆-5'), 200 μ M of dXTTP/dXATP or TTP/dATP, 100 nM of DNA polymerase in 1 \times reaction buffer (human DNA polymerase beta: 50 mM Tris-HCl (pH 8.0), 20 mM NaCl, 20 mM KCl, 10 mM MgCl₂, 2 mM DTT, 200 μ g/mL BSA, 1% [v/v] glycerol; Terminator and 9°North exo(-) DNA polymerases: 20 mM Tris-HCl (pH 8.8), 10 mM (NH₄)₂SO₄, 10 mM KCl, 2 mM MgSO₄, 0.1% Triton X-100; Dpo4 DNA polymerase: 50 mM Tris-HCl (pH 8.0), 10 mM MgCl₂, 20 mM KCl, 2 mM DTT, 100 μ g/mL BSA, 10% glycerol; KlenTaq and Taq DNA polymerase: 50 mM Tris-HCl (pH 9.2), 16 mM (NH₄)₂SO₄, 2.5 mM MgCl₂, 0.1% [v/v] Tween20). Primer was labelled by using [α -³²P]ATP according to standard techniques. Reaction mixtures were incubated for 1 h at adequate temperatures (human DNA polymerase beta and Dpo4 DNA polymerase: 37 °C; Terminator and 9°North exo(-) DNA polymerase: 55 °C; KlenTaq and Taq DNA polymerase: 72 °C). Reactions were stopped by addition of 40 μ L stop solution (80% (v/v) formamide, EDTA (20 mM), 0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol) and analyzed by 12% denaturing PAGE. Visualization was performed by phosphorimaging.

Acknowledgements

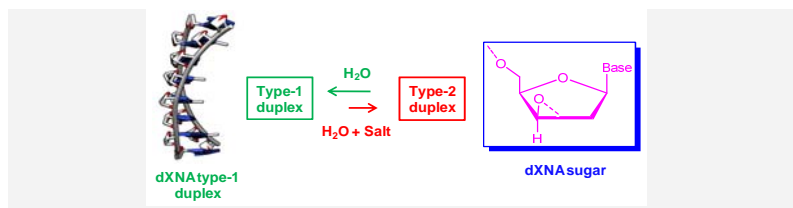
This work was financed by K. U. Leuven grants (IDO and GOA) and partially by the DFG. We are grateful to Dr. Koen Nauwelaerts for helpful discussions. We thank Prof. Arthur Van Aerschot and Guy Schepers for oligonucleotide synthesis and for providing UV-melting data, and Chantal Biernaux for editorial help.

- [1] A. Eschenmoser, *Science* **1999**, *284*, 2118-2124.
- [2] M. Beier, F. Reck, T. Wagner, R. Krishnamurthy, A. Eschenmoser, *Science* **1999**, *283*, 699-703.
- [3] K. U. Schoning, P. Scholz, S. Guntha, X. Wu, R. Krishnamurthy, A. Eschenmoser, *Science* **2000**, *290*, 1347-1351.
- [4] M. O. Ebert, B. Jaun, *Chem. Biodivers.* **2010**, *7*, 2103-2128.
- [5] P. Herdewijn, P. Marliere, *Chem. Biodivers.* **2009**, *6*, 791-808.
- [6] M. L. Stephenson, P. C. Zamecnik, *Proc. Natl. Acad. Sci. USA* **1978**, *75*, 285-288.
- [7] C. J. Wilds, Z. Wawrzak, R. Krishnamurthy, A. Eschenmoser, M. Egli, *J. Am. Chem. Soc.* **2002**, *124*, 13716-13721.
- [8] M. Egholm, O. Buchardt, L. Christensen, C. Behrens, S. M. Freier, D. A. Driver, R. H. Berg, S. K. Kim, B. Norden, P. E. Nielsen, *Nature* **1993**, *365*, 566-568.
- [9] K. Augustyns, F. Vandendriessche, A. Van Aerschot, R. Busson, C. Urbanke, P. Herdewijn, *Nucleic Acids Res.* **1992**, *20*, 4711-4716.
- [10] S. Ilin, I. Schlonvogt, M. O. Ebert, B. Jaun, H. Schwalbe, *Chembiochem* **2002**, *3*, 93-99.
- [11] F. Seela, M. Heckel, H. Rosemeyer, *Helv. Chim. Acta* **1996**, *79*, 1451-1461.
- [12] A. Ramaswamy, M. Froeyen, P. Herdewijn, A. Ceulemans, *J. Am. Chem. Soc.* **2010**, *132*, 587-595.
- [13] N. E. Poopeiko, M. Juhl, B. Vester, M. D. Sorensen, J. Wengel, *Bioorg. Med. Chem. Lett.* **2003**, *13*, 2285-2290.
- [14] B. R. Babu, Raunak, N. E. Poopeiko, M. Juhl, A. D. Bond, V. S. Parmar, J. Wengel, *Eur. J. Org. Chem.* **2005**, 2297-2321.
- [15] a) H. Rosemeyer, F. Seela, *Helv. Chim. Acta* **1991**, *74*, 748-760; b) H. Rosemeyer, M. Krecmerova, F. Seela, *Helv. Chim. Acta* **1991**, *74*, 2054-2067.
- [16] H. Rosemeyer, I. Strodtholz, F. Seela, *Bioorg. Med. Chem. Lett.* **1992**, *2*, 1201-1206.
- [17] F. Seela, K. Wormer, H. Rosemeyer, *Helv. Chim. Acta* **1994**, *77*, 883-896.
- [18] H. Rosemeyer, F. Seela, *Nucleosides & Nucleotides* **1995**, *14*, 1041-1045.
- [19] A. Schoppe, H. J. Hinz, H. Rosemeyer, F. Seela, *Eur. J. Biochem.* **1996**, *239*, 33-41.
- [20] B. F. Eichman, G. P. Schroth, B. E. Basham, P. S. Ho, *Nucleic Acids Res.* **1999**, *27*, 543-550.
- [21] B. Allart, K. Khan, H. Rosemeyer, G. Schepers, C. Hendrix, K. Rothenbacher, F. Seela, A. Van Aerschot, P. Herdewijn, *Chem. Eur. J.* **1999**, *5*, 2424-2431.
- [22] B. Furtig, C. Richter, J. Wohnert, H. Schwalbe, *Chembiochem* **2003**, *4*, 936-962.
- [23] S. S. Wijmenga, Mooren, M. M. W., & Hilbers, C. W., in *NMR of macromolecules a practical approach* (Ed.: G. C. K. Roberts), Oxford University Press, New York, **1993**, pp. 217-288.
- [24] S. S. Wijmenga, B. N. M. van Buuren, *Prog. Nucl. Mag. Res. Sp.* **1998**, *32*, 287-387.
- [25] C. D. Schwieters, J. J. Kuszewski, N. Tjandra, G. M. Clore, *J. Magn. Reson.* **2003**, *160*, 65-73.
- [26] J. A. Cromsig, C. W. Hilbers, S. S. Wijmenga, *J. Biomol. NMR* **2001**, *21*, 11-29.
- [27] K. Nauwelaerts, E. Lescrinier, P. Herdewijn, *Chem. Eur. J.* **2007**, *13*, 90-98.
- [28] E. Meggers, L. Zhang, *Acc. Chem. Res.* **2010**, *43*, 1092-1102.
- [29] M. K. Schlegel, L. O. Essen, E. Meggers, *J. Am. Chem. Soc.* **2008**, *130*, 8158-8159.
- [30] P. S. Pallan, P. Lubini, M. Bolli, M. Egli, *Nucleic Acids Res.* **2007**, *35*, 6611-6624.
- [31] S. Giesecking, K. Bergen, F. Di Pasquale, K. Diederichs, W. Welte, A. Marx, *J. Biol. Chem.* **2011**, *286*, 4011-4020.
- [32] N. Staiger, A. Marx, *Chembiochem* **2010**, *11*, 1963-1966.
- [33] S. Obeid, A. Schnur, C. Gloeckner, N. Blatter, W. Welte, K. Diederichs, A. Marx, *Chembiochem* **2011**, *12*, 1574-1580.
- [34] R. Kranaster, M. Drum, N. Engel, M. Weidmann, F. T. Hufert, A. Marx, *Biotechnol. J.* **2010**, *5*, 224-231.
- [35] D. Muller, S. Pitsch, A. Kittaka, E. Wagner, C. E. Wintner, A. Eschenmoser, *Helv. Chim. Acta* **1990**, *73*, 1410-1468.
- [36] R. Keller, *The Computer Aided Resonance Assignment Tutorial*, 1st ed., Cantina, Goldau, **2004**.
- [37] G. Ravishanker, S. Swaminathan, D. L. Beveridge, R. Lavery, H. Sklenar, *J. Biomol. Struct. Dyn.* **1989**, *6*, 669-699.
- [38] M. Krecmerova, F. Seela, *Nucleosides & Nucleotides* **1992**, *11*, 1393-1409.
- [39] J. Ludwig, *Acta Biochim. Biophys. Acad. Sci. Hung.* **1981**, *16*, 131-133.
- [40] J. L. Kosa, J. B. Sweasy, *J. Biol. Chem.* **1999**, *274*, 3851-3858.
- [41] K. A. Fiala, Z. Suo, *Biochemistry* **2004**, *43*, 2116-2125.
- [42] M. W. Southworth, H. Kong, R. B. Kucera, J. Ware, H. W. Jannasch, F. B. Perler, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 5281-5285.
- [43] K. B. Sauter, A. Marx, *Angew. Chem. Int. Ed. Engl.* **2006**, *45*, 7633-7635.
- [44] a) D. R. Engelke, A. Krikos, M. E. Bruck, D. Ginsburg, *Anal. Biochem.* **1990**, *191*, 396-400; b) D. Summerer, N. Z. Rudinger, I. Detmer, A. Marx, *Angew. Chem. Int. Ed. Engl.* **2005**, *44*, 4712-4715.
- [45] R. Lavery, & Zakrzewska, K., in *Nucleic Acid Structures* (Ed.: S. Neidle), Oxford University Press, New York, **1999**, p. 64.

Entry for the Table of Contents

Mohitosh Maiti, Vanessa Siegmund,
Mikhail Abramov, Eveline
Lesclinier, Helmut Rosemeyer,
Mathy Froeyen, Amutha
Ramaswamy, Arnout Ceulemans,
Andreas Marx, Piet Herdewijn*
..... Page – Page

**Solution Structure and
Conformational Dynamics of
Deoxyxylonucleic Acid (dXNA):
An Orthogonal Nucleic Acid
Candidate**



Orthogonal nucleic acid - Synthetic biology shows growing interest in alternate nucleic acids that can be used to safely transfer genetic information without interfering with the natural system. We have explored the structural and biochemical aspects of

xylose-DNA (dXNA) towards this goal. The study on one hand identifies properties for developing dXNA as an alternate information system, while on the other hand finds explanations why dXNA was not selected as natural genetic system.

Supporting Information

Solution Structure and Conformational Dynamics of Deoxyxylonucleic Acid (dXNA): An Orthogonal Nucleic Acid Candidate

Mohitosh Maiti^{[a],[f]}, Vanessa Siegmund^{[b],[f]}, Mikhail Abramov^[a], Eveline Lescrinier^[a], Helmut Rosemeyer^[c], Mathy Froeyen^[a], Amutha Ramaswamy^[d], Arnout Ceulemans^[e], Andreas Marx^[b], Piet Herdewijn^{[a]*}

^[a] Laboratory for Medicinal Chemistry, Rega Institute for Medical Research, Katholieke Universiteit Leuven, Minderbroederstraat 10, B-3000 Leuven, Belgium ^[b] Department of Chemistry and Konstanz Research School of Chemical Biology, University of Konstanz, Universitätsstrasse 10, 78457 Konstanz, Germany ^[c] Organic Materials Chemistry and Bioorganic Chemistry, Institute of Chemistry, University of Osnabrück, Barbarastrasse 7, D-49069 Osnabrück, Germany ^[d] Centre for Bioinformatics, Pondicherry University, Puducherry 605014, India ^[e] INPAC institute for Nanoscale Physics and Chemistry and Quantum Chemistry Group, K. U. Leuven, Celestijnenlaan 200F, 3001 Leuven, Belgium ^[f] These authors contributed equally to this work

*Corresponding author. E-mail: Piet.Herdewijn@rega.kuleuven.be

CONTENTS

Experimental Section	(Page S2-S4)
Supplementary References	(Page S4)
Figures	(Page S5-S18)
Tables	(Page S19-S27)

Experimental Section:

Oligonucleotide synthesis: Protected deoxyxylose phosphoramidite monomer (A/T/G/C) building blocks were synthesized following previously reported methods^[1]. Oligonucleotide synthesis was performed with an Expedite DNA synthesizer (Applied Biosystems) by using standard phosphoramidite approach. The oligomers were deprotected and cleaved from the solid support by treatment with methylamine (40% in water) and concentrated aqueous ammonia (1:1, 30 °C). Samples were lyophilized, then the crude mixture was analyzed by using a Mono-Q HR 5/5 anion exchange column, after which main purification was achieved by using a Mono-Q HR 10/10 (Pharmacia) anion exchange column with the following gradient system (A= pH 7.4 20 mM Tris-HCl buffer containing 10 mM NaClO₄ and 15% Acetonitrile; B= pH 7.4 20 mM Tris-HCl buffer containing 0.6 M NaClO₄ and 15% Acetonitrile). The low-pressure liquid chromatography system consisted of a Merck-Hitachi L-6200A intelligent pump, a Mono-Q HR 10/10 column (Pharmacia), an Uvicord SII 2138 UV detector (Pharmacia LKB) and a recorder. The product-containing fraction was finally desalted on a BIOGEL-P2 column (200-400 mesh) and purified aqueous samples were lyophilized. Oligonucleotides were characterized by RP-HPLC-Mass (ESI-MS) spectroscopic analysis. Calculated mass of dXNA1 and dXNA2 2488.4, found 2490.5 and 2490.6 respectively.

UV melting experiments: dXNA and corresponding DNA oligonucleotides were dissolved in a buffer solution containing NaCl (0.1 M), potassium phosphate (0.02 M, pH 7.5), and EDTA (0.1 mM). The concentration was determined by measuring the absorbance in MilliQ water at 260 nm at 80 °C, and by assuming that deoxyxylo nucleosides have the same extinction coefficients per base moiety in the denatured state as the natural nucleoside (A*, $\epsilon=15060$; T*, $\epsilon=8560$; G*, $\epsilon=12180$; C*, $\epsilon=7100$ L mol⁻¹cm⁻¹). The strand concentration was 8 μ M in all experiments. Melting curves were determined with a Varian Cary 100 BIO spectrometer. Cuvettes (1 cm path length) were maintained at constant temperature by water circulation through the cuvette holder. The temperature of the solution was measured with a thermistor that was directly immersed in the cuvette. Temperature control and data acquisition were carried out automatically with an IBM-compatible computer by using Cary WinUV thermal application software. A quick heating and cooling cycle were carried out to allow proper annealing of strands. The samples were then heated from 15 to 85 °C at a rate of 0.2 °C/min, and were cooled again at the same speed. Melting temperatures were determined by plotting the first derivative of the absorbance (at 260 nm) as a function of temperature; data plotted were the average of two runs. Up and down curves showed identical T_m values.

Circular dichroism (CD) experiments: CD spectra were measured at 10, 20,...80 °C with a Jasco 600 (Jasco, Tokyo, Japan) spectropolarimeter in thermostatically controlled 1 cm cuvette connected with a Lauda Rcs 6 bath. The oligomers were dissolved and analyzed in water and in

buffer containing sodium cacodylate (10 mM), magnesium chloride (10 mM), NaCl (100 mM), pH 7.0, with 12 μ M concentration of oligonucleotide.

NMR sample preparation: Lyophilized oligonucleotides were dissolved in D₂O (1.6 mM, strand concentration) or in H₂O-D₂O (9:1). Samples were annealed prior to NMR experiments by briefly heating at 70 °C and cooling at room temperature. The pH was adjusted by adding small amounts of 0.1 M HCl in D₂O. To investigate the effect of salt, samples in D₂O were titrated with concentrated solution of NaCl (4.8 M)/MgCl₂ (0.6 M) in D₂O and spectral changes were followed by recording a series of proton and phosphorus 1D spectra.

NMR experiments: Phosphorus (³¹P), ¹H-³¹P HETCOR, and ³¹P coupled/decoupled DQF-COSY spectra were recorded on a Bruker Avance 500 MHz spectrometer equipped with a TXI-HCP Z gradient probe. Other spectra were obtained on a Bruker Avance II 600 MHz spectrometer equipped with a 5 mm TXI-HCN Z gradient cryoprobe (BioMacs). Generally, all spectra were recorded at 10 °C for structural data; some 2D spectra were also recorded at 25 °C. Temperature dependent 1D proton spectra (in D₂O/ H₂O-D₂O) and phosphorus spectra (in D₂O) were acquired at 5, 10, 20, ...70 °C. All spectra were processed by using Bruker Topspin 2.1 software.

The 1D proton spectra in H₂O-D₂O were recorded by using excitation sculpting with gradients. The 2D NOESY experiments in H₂O-D₂O ($t_{\text{mix}}=150, 300$ ms) were recorded by using the WATERGATE water-suppression with a sweep width of 13220 Hz in both dimensions, 64 scans, 4096 data points in t_2 and 1024 FIDs in t_1 .

The 2D DQF-COSY, TOCSY and NOESY^[2] spectra in D₂O were recorded with a sweep width of 6010 Hz in both dimensions. The residual HOD peak was suppressed by low power, on resonance presaturation. Both ³¹P decoupled (on resonance, continuous decoupling) and ³¹P coupled DQF-COSY spectra were recorded under same condition. All spectra were acquired with 64 scans, 2048 data points in t_2 and 512 FIDs in t_1 . The data were apodized with a shifted sine-bell square function in both dimensions and processed to a 2k×1k matrix. For the TOCSY experiment, a clean MLEV17 pulse program was used, with mixing time of 60 ms. NOESY experiments were recorded with mixing times of 50, 100, 150 and 300 ms.

¹H-³¹P HETCOR^[3] spectra were acquired with 256 scans, 2048 data points in the proton dimension (t_2), and 256 increments in the phosphorus dimension (t_1) over sweep widths of 3750 and 808 Hz, respectively.

Natural abundance ¹H-¹³C HSQC spectra were recorded with sensitivity enhancement and gradient coherence selection optimized for selection of CH groups. To measure ¹H-¹³C correlations of the sugar part and H5-C5 correlations of cytosine, J_{CH} was set to 150 Hz and 128/2048 complex data points and 130/10 ppm spectral widths in t_1 and t_2 were used respectively. For each sample, a separate ¹H-¹³C HSQC spectrum was recorded to determine the remaining ¹H-¹³C correlations in nucleobases. It was optimized for selection of aromatic C-H

groups ($^1J_{\text{CH}}=200$ Hz) by using 128 scans and 128/2048 complex data points and 60/10 ppm spectral widths in t_1 and t_2 , respectively.

Natural abundance ^1H - ^{13}C HMBC spectra were measured by using 128 scans and 4096/128 complex data points and 10/70 ppm spectral widths in t_2 and t_1 , respectively. Delays were optimized for transfer in nucleobases by using $^1J_{\text{CH}}=200$ Hz in the low-pass filter and $^3J_{\text{CH}}=8$ Hz for long-range correlations.

Diffusion Ordered Spectroscopy (DOSY)^[4] experiment with a stimulated echo and bipolar gradient pulses was used to determine the translational diffusion constant of type-1 and type-2 duplex forms in D_2O solution with high salt concentration at 10 °C.

Molecular dynamics (MD) simulations: Amber version 9.0 was used for all simulations. The starting model duplex structures (**Table S5**) were prepared as described previously^[5]. Simulation conditions were similar for all experiments. First, Deoxyxylose-DNA was immersed in truncated octahedral box with explicit water molecules, and then sodium counter ions were added to get electrostatic neutrality of the system. In the simulation, first solvent density equilibration was performed, which was followed by production MD runs. Periodic boundary conditions, Particle Mesh Ewald summation, with shake restraints on all bonds involving H atoms were implemented in the calculations (cutoff= 10 Å, time step= 0.002 ps, heating to 283K or 300K).

References:

- [1] a) F. Seela, M. Heckel, H. Rosemeyer, *Helv. Chim. Acta* **1996**, *79*, 1451-1461; b) H. Rosemeyer, F. Seela, *Helv. Chim. Acta* **1991**, *74*, 748-760; c) H. Rosemeyer, M. Krecmerova, F. Seela, *Helv. Chim. Acta* **1991**, *74*, 2054-2067; d) F. Seela, K. Wörner, H. Rosemeyer, *Helv. Chim. Acta* **1994**, *77*, 883-896; e) K. Muhlegger, H. Von Der Eltz, F. Seela, H. Rosemeyer, Roche Diagnostics GmbH, Mannheim (DE): US patent 6,329,346 B1, **2001**.
- [2] B. Borah, J. S. Cohen, A. Bax, *Biopolymers* **1985**, *24*, 747-765.
- [3] V. Sklenar, H. Miyashiro, G. Zon, H. T. Miles, A. Bax, *Febs Lett.* **1986**, *208*, 94-98.
- [4] C. S. Johnson, *Prog. Nucl. Mag. Res. Sp.* **1999**, *34*, 203-256.
- [5] A. Ramaswamy, M. Froeyen, P. Herdewijn, A. Ceulemans, *J. Am. Chem. Soc.* **2010**, *132*, 587-595.

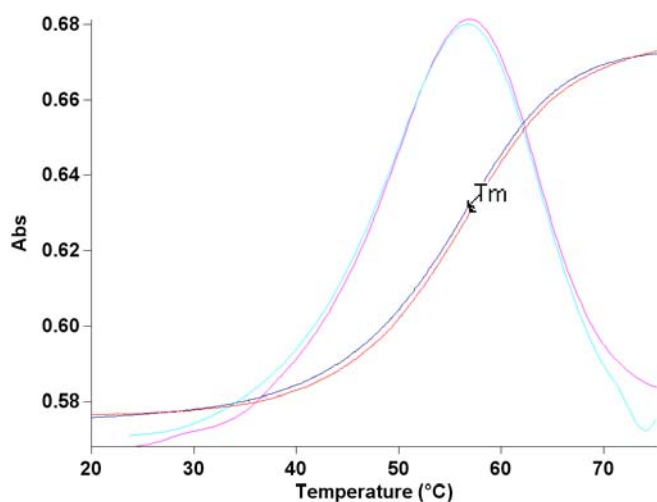


Figure S1: Sigmoidal (thermal) UV-melting curves and the corresponding first-derivative curves for T_m analysis (Absorbance (Abs) at 260 nm wavelength) of the hydrogen-bonded form of dXNA1 oligonucleotide in aqueous buffer solution. Both the up (20 to 75 °C) and down (75 to 20 °C) curves are indicative for a cooperative melting profile.

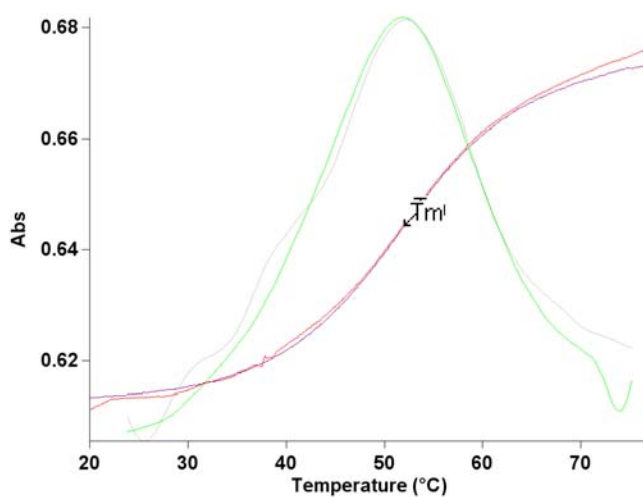


Figure S2: Sigmoidal (thermal) UV-melting curves and the corresponding first-derivative curves for T_m analysis (Absorbance (Abs) at 260 nm wavelength) of the hydrogen-bonded form of dXNA2 oligonucleotide in aqueous buffer solution. Both the up (20 to 75 °C) and down (75 to 20 °C) curves are indicative for a cooperative melting profile.

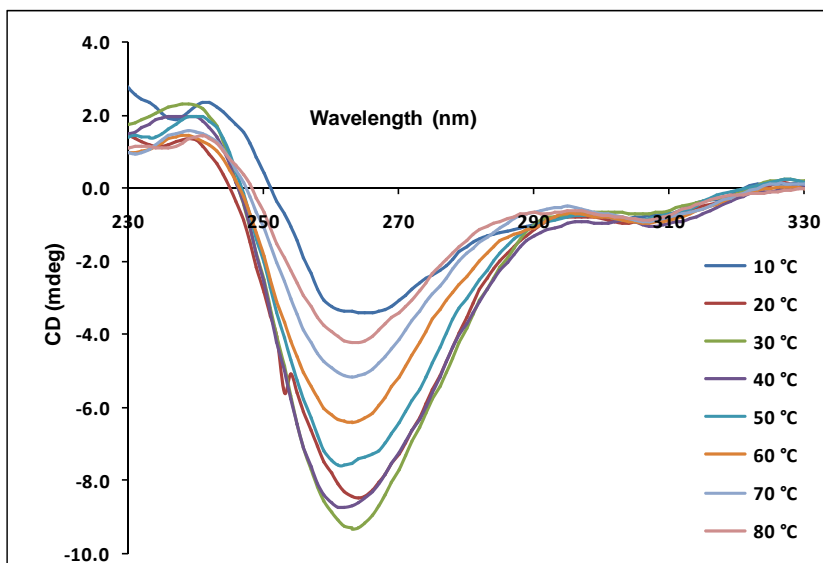


Figure S3: Temperature dependent CD spectra of dXNA2 oligonucleotide in H₂O (12 μM) at pH 7.0.

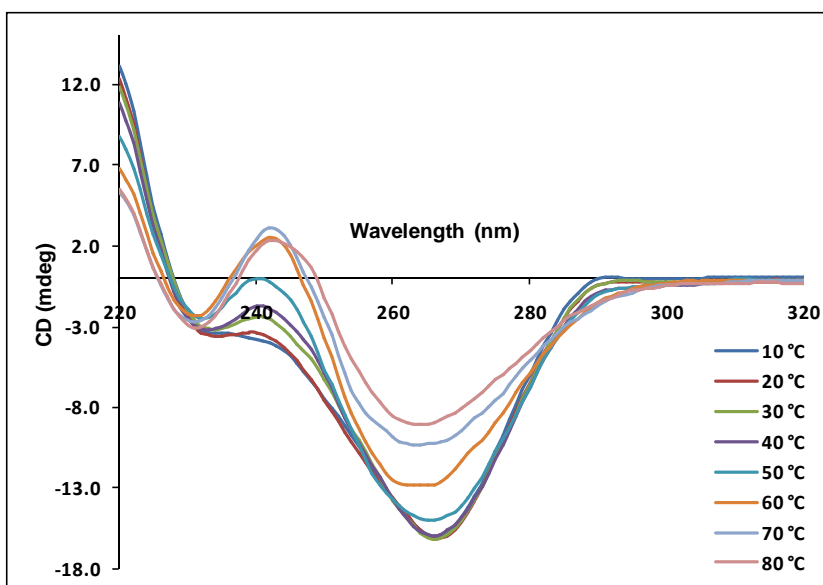


Figure S4: Temperature dependent CD spectra of dXNA2 oligonucleotide (12 μM) in aqueous buffer solution [containing sodium cacodylate (10 mM), NaCl (100 mM), MgCl₂ (10 mM)] of pH 7.0.

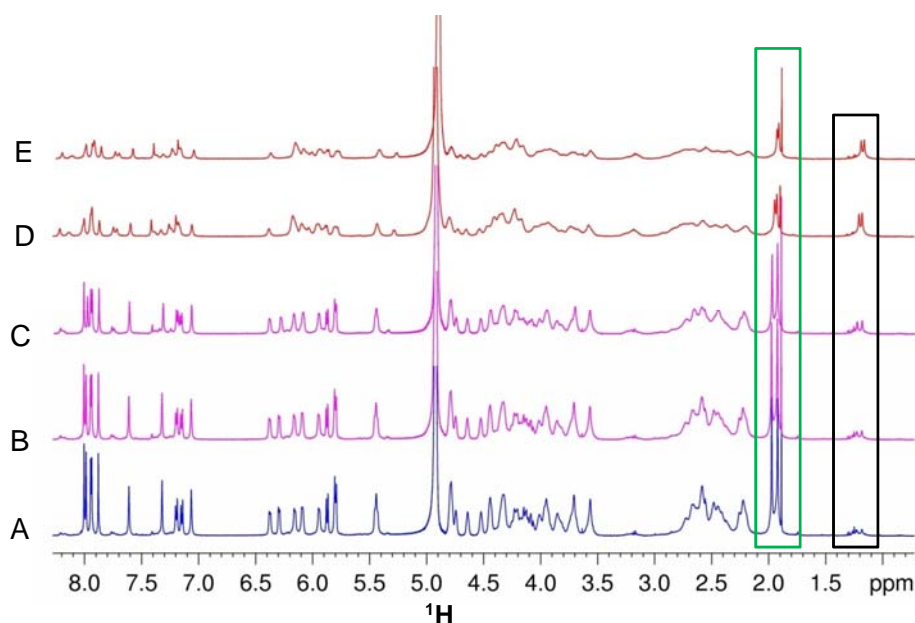


Figure S5: Stack plot of 1D proton spectra of dXNA1 in D₂O at 10 °C (pD 7.0) with different salt concentrations, (A) without salt, (B) 16 mM NaCl, (C) 78 mM NaCl, (D) 77 mM NaCl and 15 mM MgCl₂, (E) 185 mM NaCl and 23 mM MgCl₂. Intensity decrease of methyl (T) protons of the type-1 duplex form in water and the simultaneous increase in the intensity of methyl (T) protons in (up-field) other type of structural form (subsequently characterized as type-2 duplex form with higher salt concentration) are highlighted by green and black box respectively.

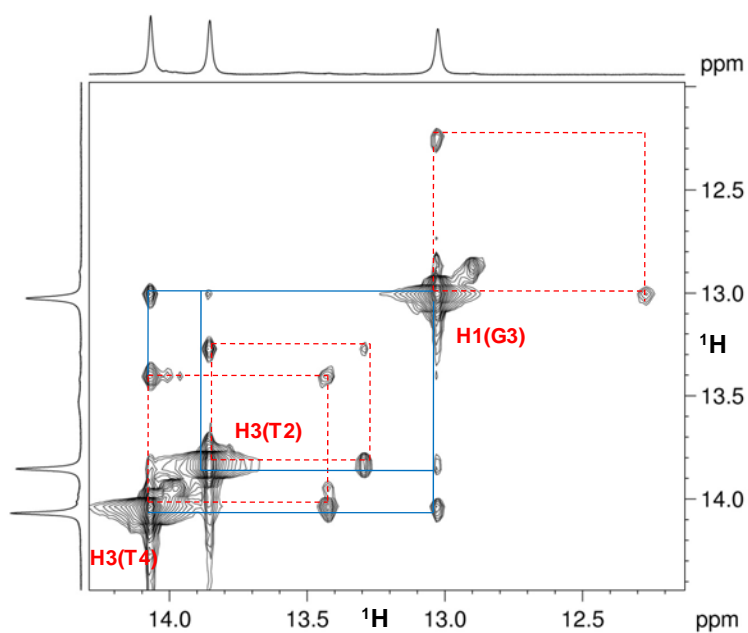


Figure S6: Hydrogen-bonded imino proton region of ^1H - ^1H 2D-NOESY (300 ms) spectrum of dXNA1 in H_2O - D_2O at 10 °C (pH 6.5). NOE correlations between imino protons of type-1 duplex are shown by blue solid lines and the cross-peaks arising from conformational exchange with type-2 duplex are indicated in red dashed lines.

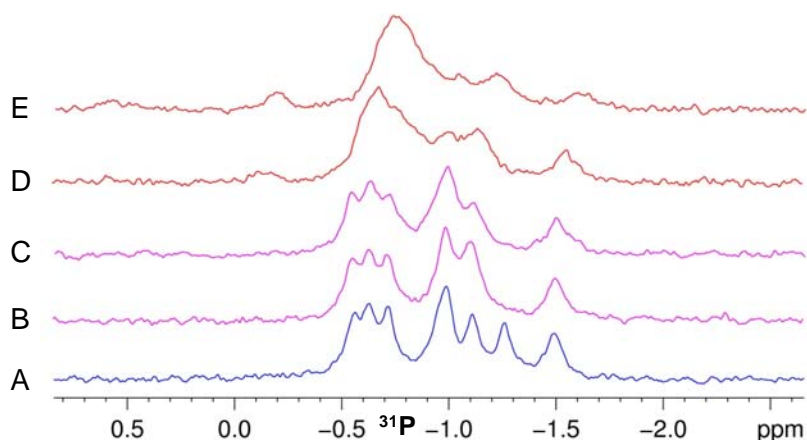


Figure S7: Stack plot of 1D phosphorus spectra of dXNA2 in D_2O at 10 °C (pD 6.5) with different salt concentration, (A) with no salt, (B) 86 mM NaCl, (C) 209 mM NaCl, (D) 207 mM NaCl and 5 mM MgCl_2 , (E) 203 mM NaCl and 15 mM MgCl_2 , showing spectral exchange broadening due to conformational equilibrium between two types of duplex forms.

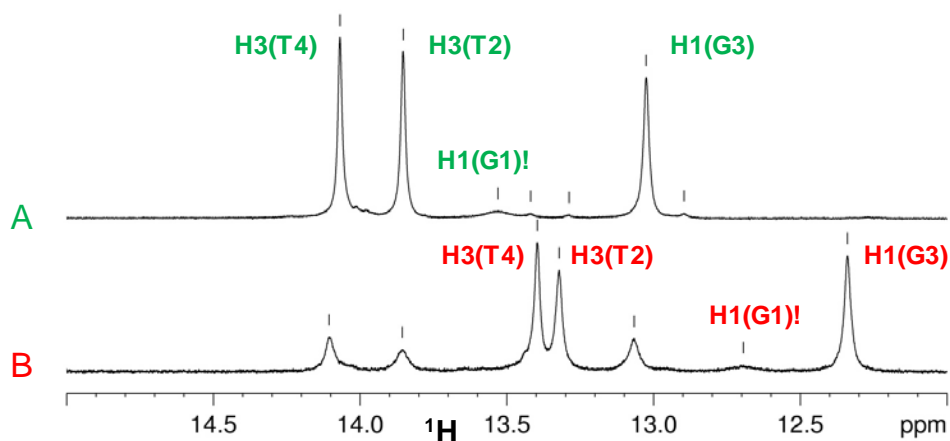


Figure S8: Stack plot of 1D hydrogen-bonded imino-proton spectra of dXNA1 in H_2O - D_2O (9:1) at 10 °C (pH 6.5), (A) without salt (type-1 duplex structure), (B) with 185 mM NaCl and 23 mM MgCl_2 (type-2 structure). H3(T) and H1 (G) imino signals are from A-T, G-C Watson-Crick base pairing steps respectively. Assignments of imino signals are based on the corresponding 2D Watergate NOESY spectra.

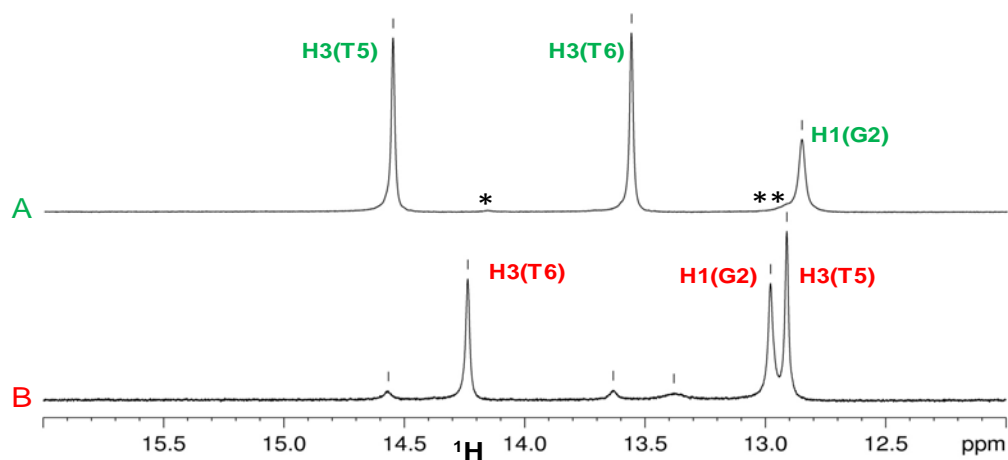


Figure S9: Stack plot of hydrogen-bonded imino-proton spectra of dXNA2 in $\text{H}_2\text{O}-\text{D}_2\text{O}$ at 10 °C (pH 6.5), (A) without salt (type-1 duplex), (B) with 203 mM NaCl and 15 mM MgCl_2 (type-2 duplex). H3(T) and H1 (G) imino signals are from A-T, G-C Watson-Crick base pairing steps respectively. Assignments of imino signals are based on the corresponding 2D Watergate NOESY spectra. * Indicates the existence of a very small peak.

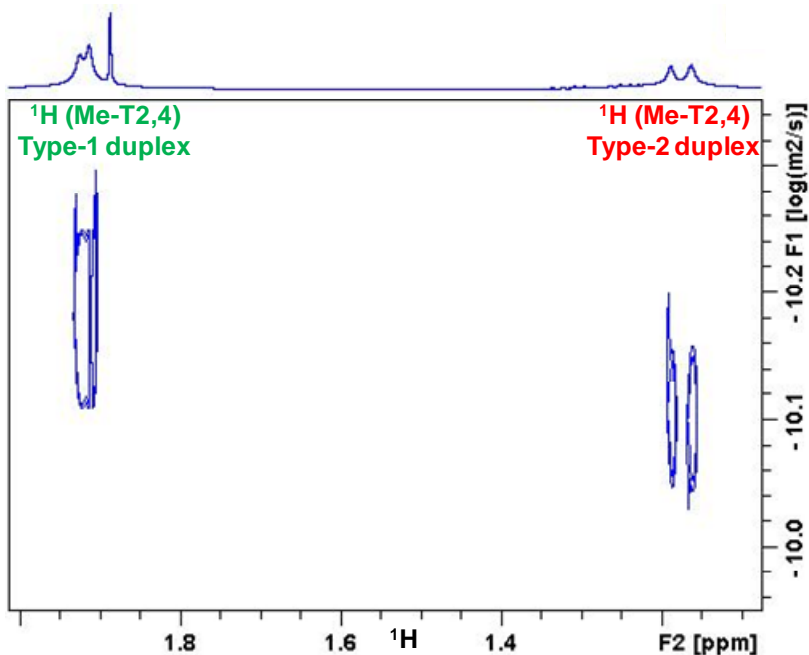


Figure S10: 2D DOSY-NMR plot of type-1 and type-2 duplex form of dXNA1 at 10 °C in D_2O with 185 mM NaCl and 23 mM MgCl_2 (pD 7.0). The plot shows faster diffusion of type-2 duplex than type-1 duplex.

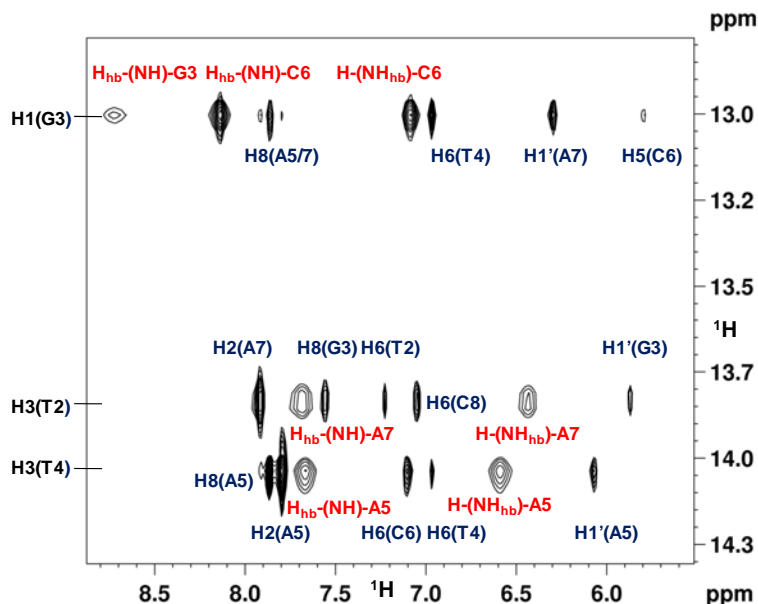


Figure S11: Expansion of 2D-NOESY (300 ms) spectrum (aromatic, anomeric and amino protons vs. hydrogen-bonded imino proton region) of dXNA1 in H₂O-D₂O at 10 °C (pH 6.5). Cross-peaks are assigned by their corresponding atom names, and residue names are within parenthesis. Hydrogen-bonded and free amino protons are denoted (in red) as H_{hb}-(NH)- and H-(NH_{hb})- respectively.

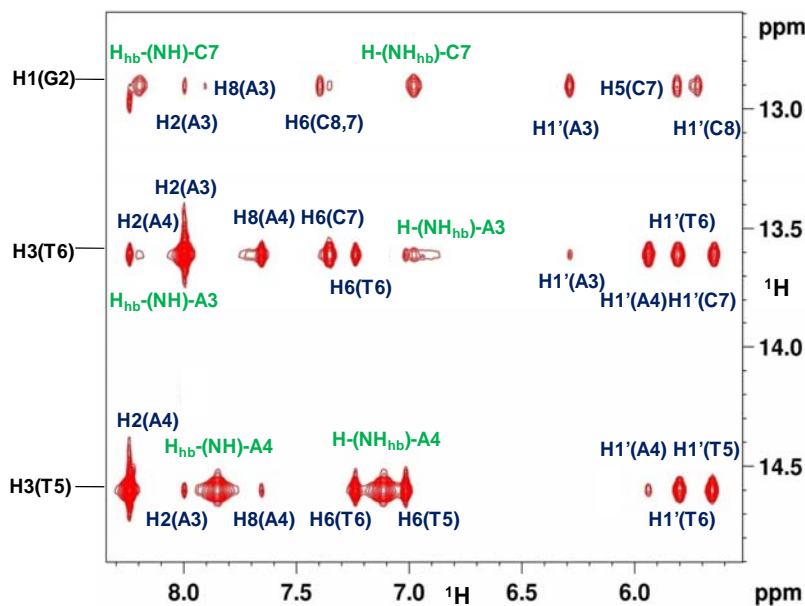


Figure S12: Expansion of 2D-NOESY (300 ms) spectrum (aromatic, anomeric and amino protons vs. hydrogen-bonded imino proton region) of dXNA2 in H₂O-D₂O at 10 °C (pH 6.5). Cross-peaks are assigned by their corresponding atom names, and residue names are within parenthesis. Hydrogen-bonded and free amino protons are denoted (in green) as H_{hb}-(NH)- and H-(NH_{hb})- respectively.

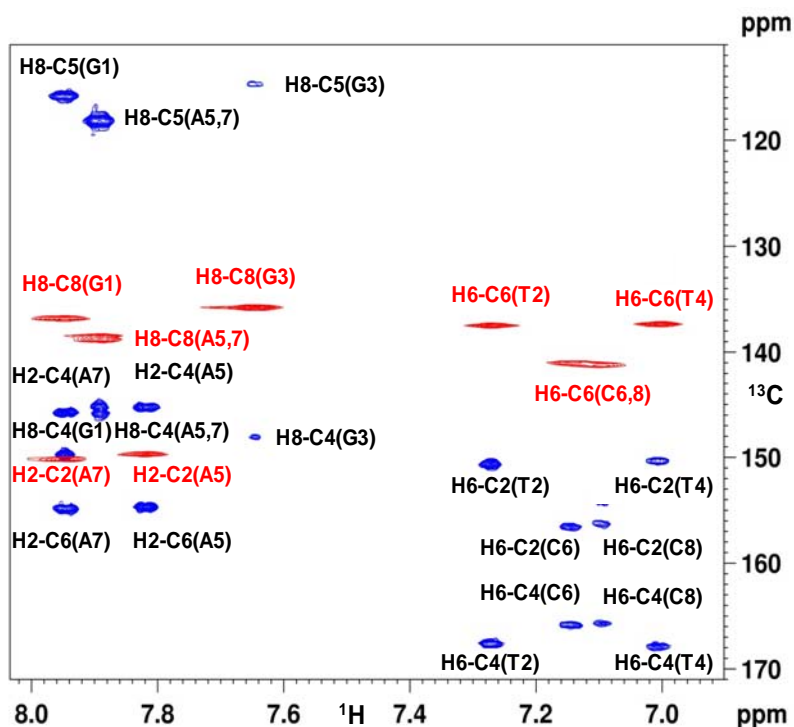


Figure S13: Overlay of ¹H-¹³C HSQC (red cross-peaks and labels) and HMBC (blue cross-peaks with black labels) spectrum in aromatic-proton/carbon region (except H5, C5 of cytosine) of dXNA1 type-1 duplex in D₂O at 10 °C. Assignments are indicated by their respective proton-carbon atom names, residue numbers are within parenthesis.

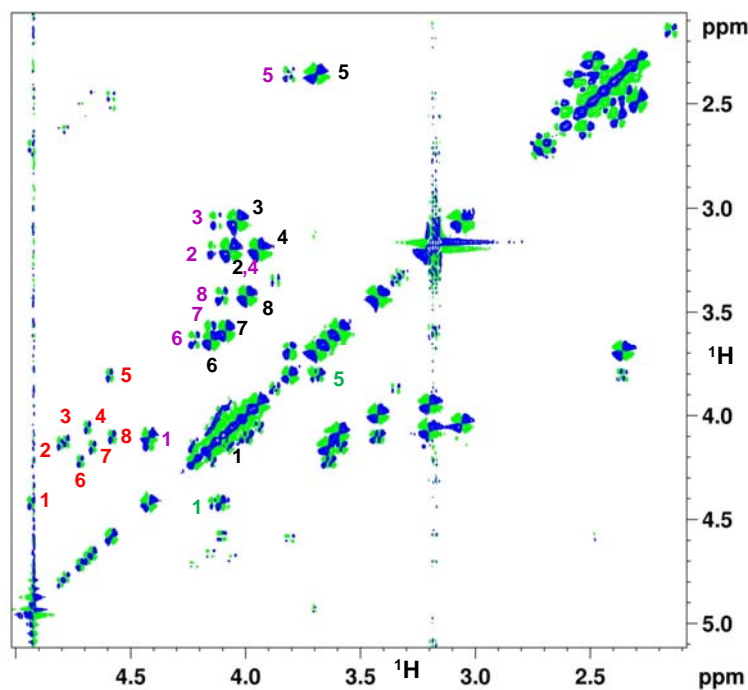


Figure S14: ^{31}P -decoupled DQF-COSY spectrum (sugar proton region, except anomeric proton) of dXNA2 duplex in D_2O (10°C). For clarity assignments are only shown by corresponding residue numbers (5' to 3'-end). $\text{H}_3'-\text{H}_4'$, $\text{H}_4'-\text{H}_5''$, $\text{H}_5'-\text{H}_5''$ and $\text{H}_5'-\text{H}_4'$ cross-peaks (only assigned in one dimension) are marked in red, purple, black and green respectively.

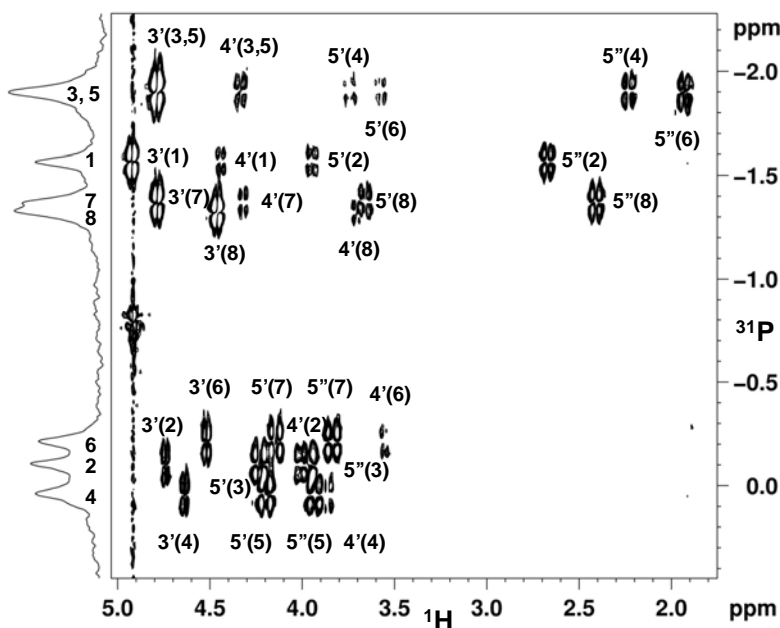


Figure S15: ^1H - ^{31}P HETCOR spectrum of dXNA1 duplex, recorded at 10 °C in D_2O . Phosphorus atoms numbered from 5' to 3' direction (with an extra ^{31}P atom at the 3'-end). All the H3', H4', H5' and H5'' protons are assigned with their corresponding residue number. Remarkably, H5'' protons of pyrimidine residues (after a preceding purine residue) are highly shielded and appeared up-field. Appearance of H4'-P3' (4J) cross-peaks is indicative for a W-conformation involving the related backbone atoms.

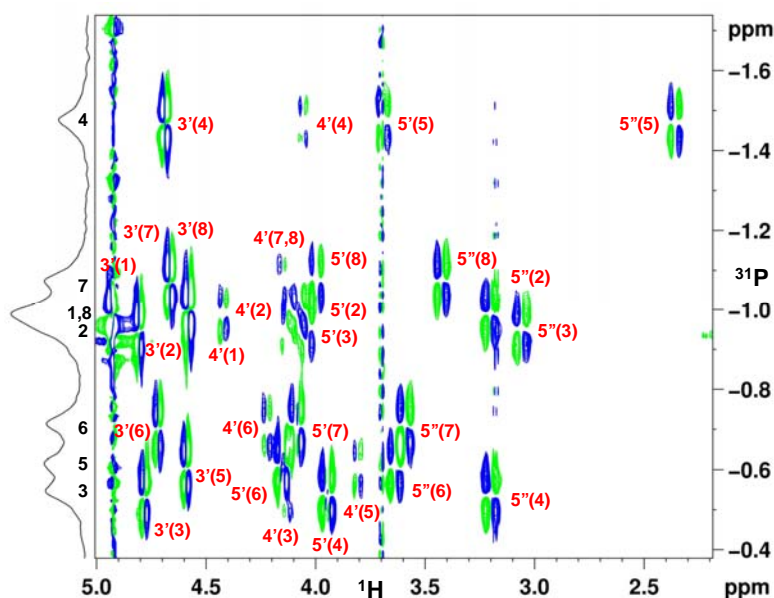


Figure S16: ^1H - ^{31}P HETCOR spectrum, recorded at 10 °C in D_2O for dXNA2 duplex. Phosphorus atoms numbered from 5' to 3' direction (with an extra ^{31}P atom at the 3'-end). All the H3', H4', H5' and H5'' protons are assigned with their corresponding residue number. Remarkably, H5'' protons of T5 (pyrimidine) residue after the preceding A4 (purine) residue are highly shielded and appeared at up-field. Appearance of H4'-P3' (4J) cross-peaks is indicative for a W-conformation involving the related backbone atoms.

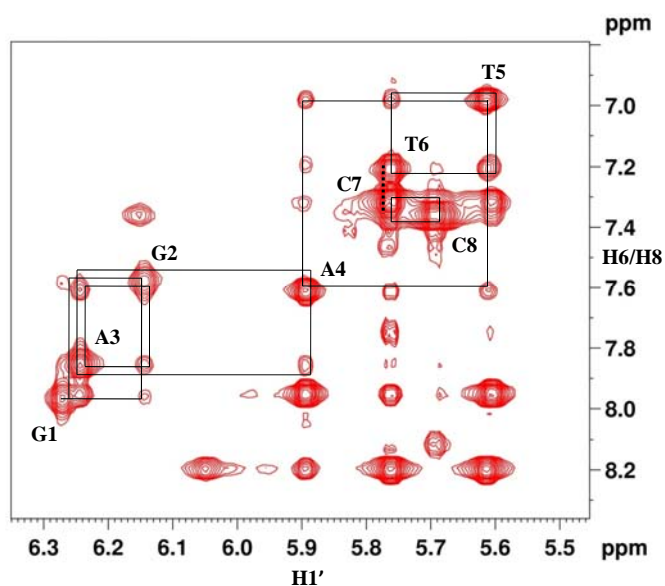


Figure S17: Expansion of NOESY spectrum in D₂O (300 ms, at 10 °C) for dXNA2 (5'-G¹GAATTCC⁸-3') duplex, showing a less pronounced NOE walk between H6/H8-H1' intra-residue and H6/H8(n)-H1'(n+1), H1'(n)-H6/H8(n+1) inter-residue protons.

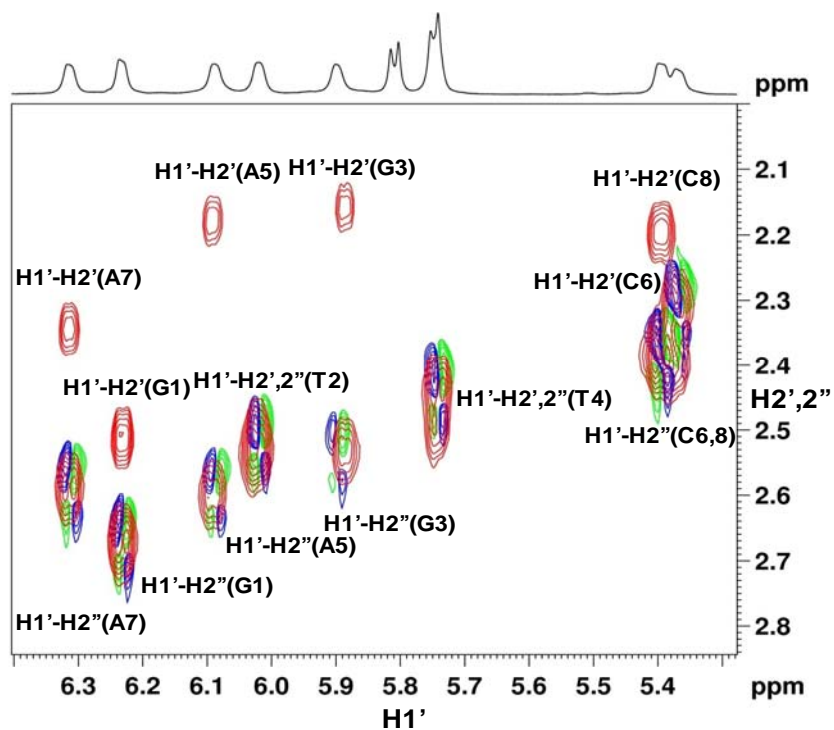


Figure S18: Overlay of ^1H - ^1H DQF-COSY and TOCSY spectrum ($\text{H1}'$ - $\text{H2}'$, 2'' proton region) of dXNA1 duplex (type-1) in D_2O at 10 °C. Residue numbers are indicated within parenthesis. Absence of $\text{H1}'$ - $\text{H2}'$ cross-peaks in DQF-COSY spectrum and the appearance of $\text{H1}'$ peak as doublet in the 1D proton (projection) spectrum ($^3J_{\text{H1}'\text{-H2}'} \approx 0$ Hz) are indicative of a predominant North (N)-type sugar pucker of the five-membered 2'-deoxyxylose sugar ring.

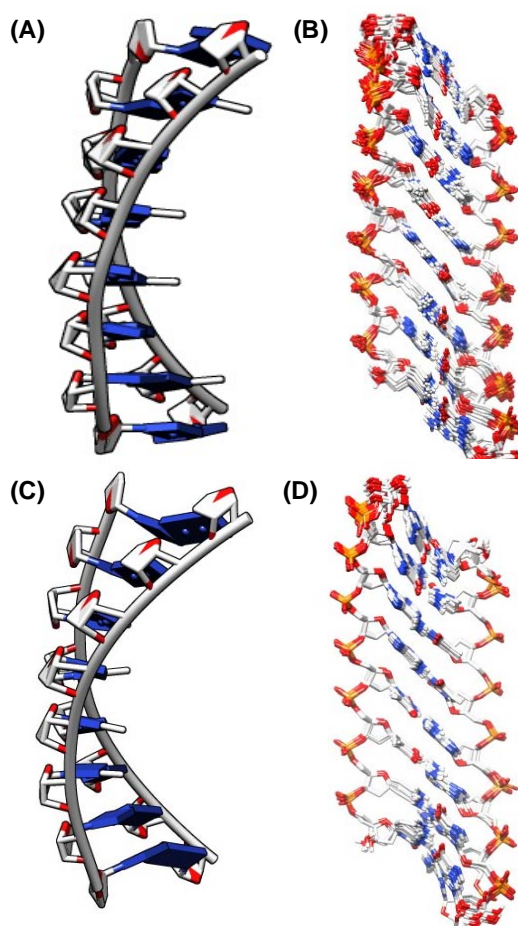


Figure S19: Solution structure of minimum energy dXNA1 (A) and dXNA2 (C) type-1 duplex in ribbon representation. Overlay (view in major groove) of 20 refined dXNA1 (B) and dXNA2 (D) duplex structures.

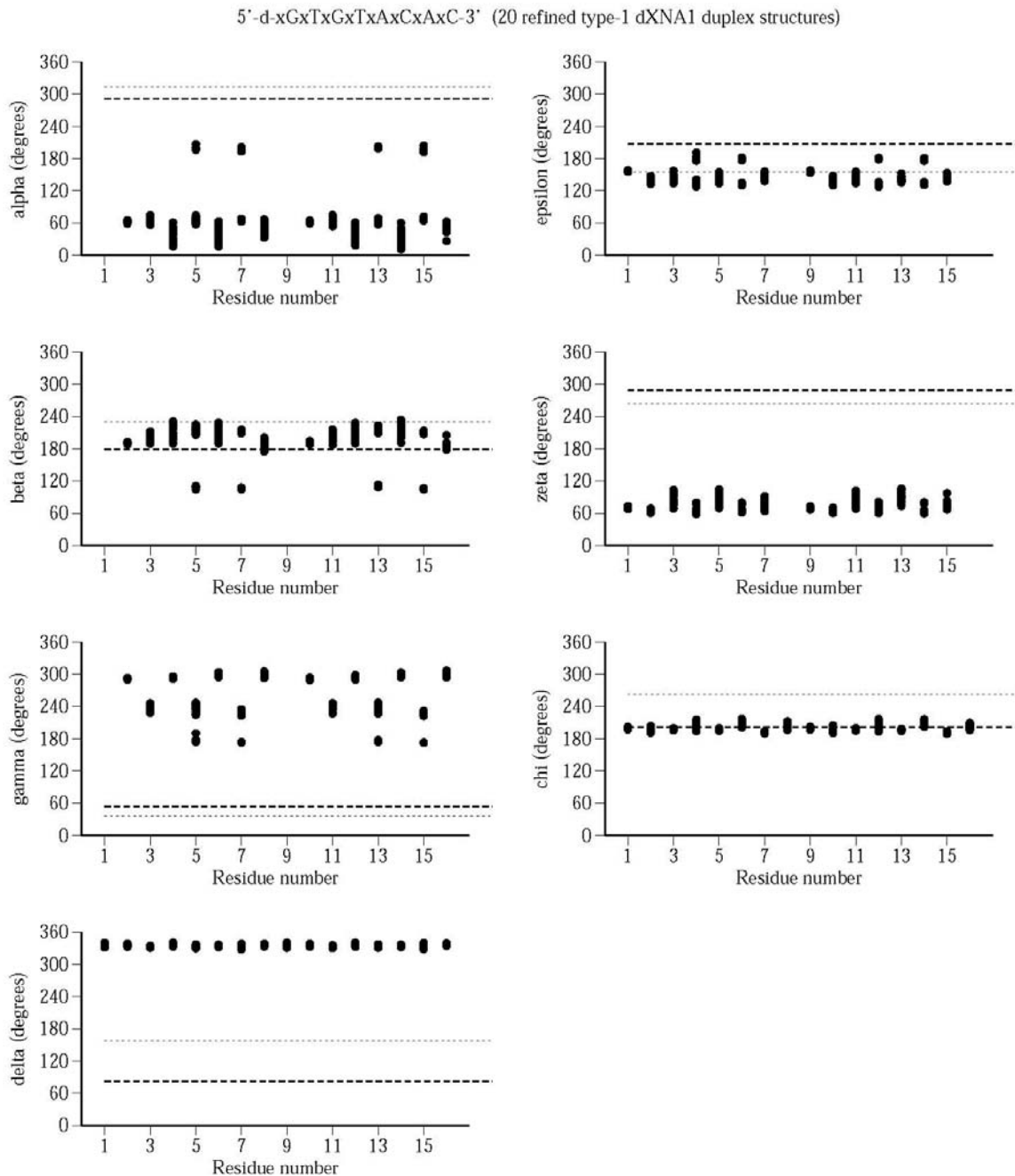


Figure S20: Variation of backbone and glycosyl torsion angles in the individual residues of the 20 final duplex (type-1) structures calculated for dXNA1 oligonucleotide. For comparison, A-type RNA and B-type DNA values are given by dashed and dotted lines, respectively.

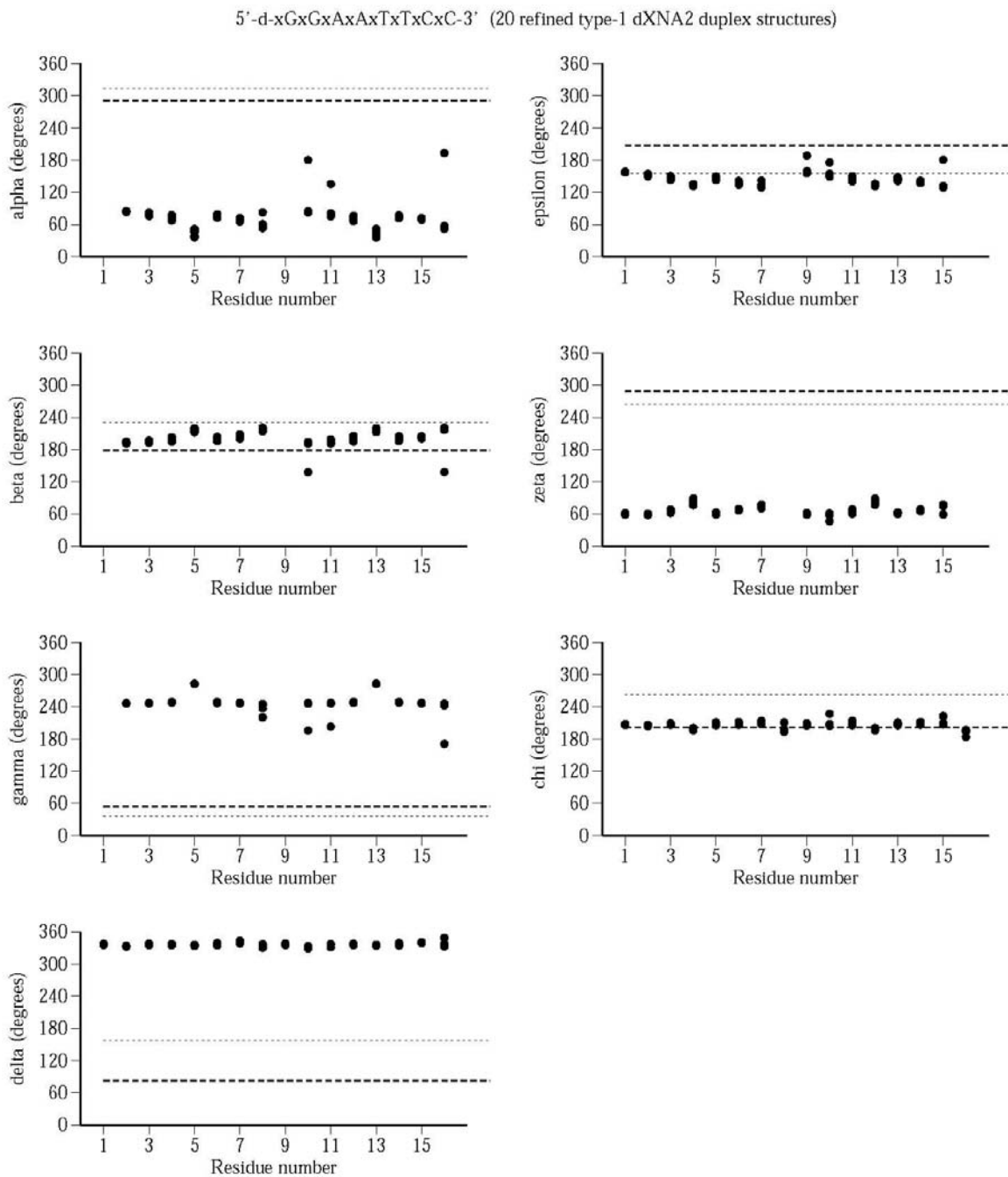


Figure S21: Variation of backbone and glycosyl torsion angles in the individual residues of the 20 final duplex (type-1) structures calculated for dXNA2 oligonucleotide. For comparison, A-type RNA and B-type DNA values are given by dashed and dotted lines, respectively.

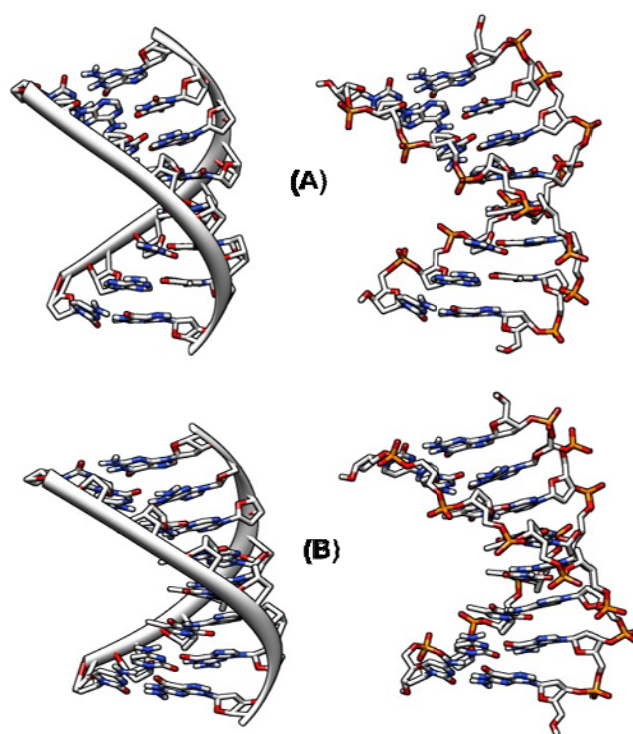


Figure S22: Left-handed structures (alternative form of an extended ladder like duplex) of (A) dXNA1 and (B) dXNA2 duplex derived from molecular dynamics simulations. Left: with backbone ribbon and Right: without ribbon.

Table S1: Coupling constants extracted from ^{31}P coupled/decoupled DQF-COSY spectra (at 10 °C in D_2O , 1.6 mM) for the duplex form of dXNA1 and dXNA2. $^3J_{\text{H}5'-\text{P}5'}/^3J_{\text{H}5''-\text{P}5'}$ couplings were estimated by measuring the ^{31}P passive coupling in $\text{H}5'-\text{H}5''/\text{H}5''-\text{H}5'$ cross-peaks in ^{31}P coupled DQF-COSY spectra (by the difference of Hz values between positive to negative lobe Hz separation in ^{31}P coupled DQF-COSY and ^{31}P decoupled DQF-COSY spectrum). $^3J_{\text{H}4'-\text{H}5'}/^3J_{\text{H}4'-\text{H}5''}$ couplings were estimated by measuring the couplings of $\text{H}4'$ proton (peak top to bottom in Hz) in the corresponding 1D slice (in row) of $\text{H}5'-\text{H}4'/\text{H}5''-\text{H}4'$ cross-peaks of ^{31}P decoupled DQF-COSY spectrum. Based on these couplings the backbone torsion angles (β , γ , ϵ) were restrained in the structure calculation of dXNA duplexes. * These coupling constants were not determined because of cross-peak overlap or overlap with the diagonal.

Residues	Couplings	dXNA1 (Hz)	dXNA2 (Hz)
1-8	$^3J_{\text{H}3'-\text{P}3'}$	3-7	4.5-8
3, 5, 7 (Pu)	$^3J_{\text{H}5'-\text{P}5'}$	8-10	
3, 5, 7 (Pu)	$^3J_{\text{H}5''-\text{P}5'}$	8-10	
2, 4, 6, 8 (PyT)	$^3J_{\text{H}5'-\text{P}5'}$	~1-3	
2, 4, 6, 8 (PyT)	$^3J_{\text{H}5''-\text{P}5'}$	~1-3	
3, 5, 7 (Pu)	$^3J_{\text{H}4'-\text{H}5'}$	~0	
3, 5, 7 (Pu)	$^3J_{\text{H}4'-\text{H}5''}$	7-8	
2, 4, 6, 8 (PyT)	$^3J_{\text{H}4'-\text{H}5'}$	11-13	
2, 4, 6, 8 (PyT)	$^3J_{\text{H}4'-\text{H}5''}$	4.5-5.0	
2, 3, 4, 6, 7, 8	$^3J_{\text{H}5'-\text{P}5'}$		4-8
2, 3, 4, 6, 7, 8	$^3J_{\text{H}5''-\text{P}5'}$		4-8
5 (T)	$^3J_{\text{H}5'-\text{P}5'}$		~3
5 (T)	$^3J_{\text{H}5''-\text{P}5'}$		~2
5 (T)	$^3J_{\text{H}4'-\text{H}5'}$		7
5 (T)	$^3J_{\text{H}4'-\text{H}5''}$		~2
3, 6, 7, 8	$^3J_{\text{H}4'-\text{H}5'}$		ND*
3, 6, 7, 8	$^3J_{\text{H}4'-\text{H}5''}$		6-7

Table S2: Structure-determination statistics on the final set of twenty refined structures calculated with all experimental restraints for type-1 dXNA1 and dXNA2 duplex. Standard deviations are provided within parenthesis.

	dXNA1	dXNA2
Total energy [kcal mol ⁻¹]	129.38 (3.83)	173.25 (1.59)
Experimental restraints		
Intranucleotide distances	96	96
Internucleotide distances	104	254
Hydrogen bond distances	20	20
Dihedral angles	90	90
RMS deviations [Å]		
From distance restraints	0.045 (0.008)	0.015 (0.003)
From dihedral restraints	0.252 (0.102)	0.989 (0.045)
From average structure for all heavy atoms	0.571 (0.295)	0.450 (0.198)
Restraint violations		
Number of distance violations (>0.5 Å)	0	0
Number of dihedral violations (>5°)	0	0

Table S3: Helical parameters for natural B-DNA and A-RNA duplexes, left-handed (L-dXNA) and intermediate (I-dXNA) structural model of dXNA duplexes derived from previously reported MD-simulation, and the NMR solution structures of two type-1 dXNA duplexes under this study. Standard values for DNA, RNA are from ref.^[45] and values for L-dXNA, I-dXNA from ref.^[12]. Average and standard deviations (\pm) were calculated for an ensemble of 20 refined NMR structures using CURVES program.

Parameters	B-DNA	A-RNA	L-dXNA	I-dXNA	dXNA1	dXNA2
Base pair-Axis Parameters						
X-displacement (Å)	-0.7	-5.3	-1.0	0	1.4 \pm 0.5	-0.3 \pm 0.1
Y-displacement (Å)	0	0	-0.1	0	0 \pm 0.3	0 \pm 0.3
Inclination (°)	-5.9	15.8	24.0	-49.0	-52.2 \pm 0.4	-49 \pm 0.5
Tip (°)	0	0	-1.0	2.5	0 \pm 0.7	0 \pm 0.5
Global Inter-Base pair Parameters						
Shift (Å)	0	0	0	0	0 \pm 0.010	0 \pm 0.007
Slide (Å)	0	0	0.3	0.3	0 \pm 0.02	0 \pm 0.02
Rise (Å)	3.4	2.8	4.0	5.0	5.8 \pm 0.06	5.1 \pm 0.05
Tilt (°)	0	0	0.3	-0.3	0.2 \pm 0.71	0 \pm 0.08
Roll (°)	0	0	13.0	20.3	8.6 \pm 1.3	12.3 \pm 1.3
Twist (°)	36.0	32.7	-20.0	-6.6	2.7 \pm 0.5	3.5 \pm 0.6

Table S4: Experimental (in D₂O at 10 °C, pD 7.0) and calculated chemical shifts (by Nuchemics program) for non-exchangeable protons of dXNA1 residues in type-1 duplex. Experimental and calculated chemical shifts of H5' and H5'' protons are highlighted by green background. Generally, for all residues H5'' protons are shielded than H5' in both experimental and calculated shifts. Experimentally, unusual high shielding effect was observed for H5'' protons of all alternate pyrimidine residues. Calculated shift values from the NMR structure also show this trend in alternate steps.

Residue	Proton	Experimental chemical shift (ppm)	Calculated chemical shift (ppm)	$\Delta(\text{Exp-Calc})$ (ppm)
G1 (5')	H1'	6.234	5.970	0.264
	H2'	2.506	2.239	0.267
	H2''	2.666	2.767	-0.101
	H3'	4.857	4.857	0.000
	H4'	4.377	4.349	0.028
	H5''	4.013	4.200	-0.187
	H5'	4.054	4.234	-0.180
	H8	7.953	8.184	-0.231
T2	H1'	6.021	5.923	0.098
	H2'	2.631	2.161	0.470
	H2''	2.524	2.593	-0.069
	H3'	4.678	4.664	0.014
	H4'	3.959	3.833	0.126
	H5''	2.629	2.391	0.238
	H5'	3.893	3.380	0.513
	H6 H(me)	7.272 1.908	7.178 2.045	0.094 -0.137
G3	H1'	5.899	6.039	-0.140
	H2'	2.17	2.038	0.132
	H2''	2.532	2.709	-0.177
	H3'	4.724	4.771	-0.047
	H4'	4.273	4.264	0.009
	H5''	3.919	3.946	-0.027
	H5'	4.169	3.857	0.312
	H8	7.648	7.726	-0.078
	H1'	5.743	5.723	0.020
	H2'	2.501	2.097	0.404
	H2''	2.414	2.488	-0.074
	H3'	4.575	4.588	-0.013

T4	H4'	3.809	3.734	0.075
	H5''	2.205	2.650	-0.445
	H5'	3.693	3.207	0.486
	H6	7.005	6.796	0.209
	H(me)	1.855	1.868	-0.013
A5	H1'	6.088	6.232	-0.144
	H2'	2.162	2.064	0.098
	H2''	2.589	2.796	-0.207
	H3'	4.728	4.761	-0.033
	H4'	4.276	4.268	0.008
	H5''	3.872	4.005	-0.133
	H5'	4.134	3.922	0.212
	H2	7.819	7.813	0.006
C6	H8	7.895	8.026	-0.131
	H1'	5.368	5.267	0.101
	H2'	2.408	1.889	0.519
	H2''	2.3	2.371	-0.071
	H3'	4.453	4.475	-0.022
	H4'	3.496	3.485	0.011
	H5''	1.85	2.720	-0.870
	H5'	3.5	3.167	0.333
	H5	5.809	6.048	-0.239
A7	H6	7.148	6.832	0.316
	H1'	6.315	5.982	0.333
	H2'	2.336	2.160	0.176
	H2''	2.581	2.712	-0.131
	H3'	4.721	4.644	0.077
	H4'	4.27	4.151	0.119
	H5''	3.784	3.868	-0.084
	H5'	4.076	3.807	0.269
	H2	7.949	7.945	0.004
C8 (3')	H8	7.893	7.733	0.160
	H1'	5.395	5.634	-0.239
	H2'	2.198	1.702	0.496
	H2''	2.373	2.393	-0.020
	H3'	4.397	4.361	0.036
	H4'	3.649	3.509	0.140
	H5''	2.341	1.887	0.454
	H5'	3.595	2.687	0.908
C8 (3')	H5	5.749	5.984	-0.235
	H6	7.096	6.635	0.461

Table S5: Simulation conditions for 8-mer dXNA duplexes. ^a Starting structures were constructed from structures of the 29-mer MD trajectory, Ref.^[12] ^b Time is production runtime. ^c Repeated simulation with cutoff = 12 instead of 10 Å.

id	dXNA duplex (8mer, anti-parallel)	Starting conformation	Simulation temperature & time ^b	End conformation	Transition point (right to left-handed) ^b
Xyl8a	5'-GTGTACAC	left-handed ^a	300K, 16.5ns	left-handed	-
Xyl8b	5'-GGAATTCC	left-handed ^a	300K, 16.5ns	left-handed	-
Xyl8c	5'-GTGTACAC	NMR model dXNA1	283K, 10.0ns	NMR like conformation	no transition
			300K, 14.5ns	NMR like conformation	no transition yet after 14.5ns
Xyl8d	5'-GTGTACAC	right-handed ^a	283K, 8.5ns	left-handed	~4ns
			300K, 4.5ns		~1.5ns
Xyl8e	5'-GGAATTCC	right-handed ^a	283K, 8.5ns	left-handed	during equilibration
			300K, 6.5ns		in equil. phase
Xyl8f	5'-GGAATTCC	NMR model dXNA2	283K, 10ns	NMR like conformation	no transition
			283K, 5ns ^c	left handed	transition at ~3.5ns
			300K, 4.5ns	left-handed	~2.0ns

Table S6: Helical parameters for the 29-mer and the 13-mer dXNA duplex structures from the simulations of Ref. ^[12]

Parameters	29mer R-dXNA	29mer transition	29mer L-dXNA	13mer start A	13mer end A I-dXNA	13mer start B	13mer/end B
Base pair-Axis Parameters							
X-displ (Å)	0.4	1.8	-1.0	0.6	0.3	0.6	1.4
Y-displ (Å)	1.3	-1.3	-0.1	-0.2	-0.3	-0.2	0.4
Inclination (°)	88.0	37.1	24.0	-29.1	-49.2	-29.1	-42.5
Tip (°)	-70.7	-3.7	-1.0	-1.7	2.5	-1.7	3.1
Global Inter-Base pair Parameters							
Shift (Å)	0.4	0.0	0	0.2	-0.1	0.2	-0.1
Slide (Å)	0.1	0.1	0.3	0.1	0.3	0.1	0.1
Rise (Å)	3.2	5.6	4.0	4.5	5.0	4,5	5.2
Tilt (°)	0.5	5.0	0.3	2.1	-0.3	2.1	-1.9
Roll (°)	7.5	14.1	13.0	20.5	20.3	20.5	22.9
Twist (°)	20.8	0.5	-20.0	8.9	-6.6	8.9	0.3
Sugar pucker	C2'endo	Between C2'endo and C3'endo	C3'endo	C3'endo	C3'endo	Between C2'endo and C3'endo	C3'endo

Table S7: Helical parameters for start and end structures of the xyl8c and xyl8f 8-mer simulations from table S5.

Parameters	Xyl8c	Xyl8c	Xyl8f	Xyl8f
	283K	300K	283K	300K
	end	end	end	end
Base pair-Axis Parameters				
X-displ (Å)	0.7	4.1	2.0	-0.9
Y-displ (Å)	-0.6	0.1	0.3	2.0
Inclination (°)	-48.4	-34.8	-39.6	-25.1
Tip (°)	-4.3	9.7	0.4	14.0
Global Inter-Base pair Parameters				
Shift (Å)	0.1	-0.2	0.0	0.4
Slide (Å)	0.3	0.2	0.0	0.1
Rise (Å)	5.2	5.7	5.1	4.0
Tilt (°)	-0.5	-1.8	-1.2	-3.9
Roll (°)	10.0	14.5	16.7	15.7
Twist (°)	-9.9	-7.6	0.7	-16.5
Pucker	C3'endo	Around C3'endo	Around C3'endo	Around C3'endo

Table S8: Helical parameters for start, transition from right to left-handed structures, and end structures of the xyl8d and xyl8e 8-mer simulations as defined in Table S5.

Parameters	Xyl8d	Xyl8d	Xyl8d	Xyl8d	Xyl8d	Xyl8e	Xyl8e	Xyl8e	Xyl8e	Xyl8e
	-	283K	283K	300K	300K	-	283K	283K	300K	300K
	start	trans	end	trans	end	start	trans	end	trans	end
Base pair-Axis Parameters										
X-displ (Å)	-0.2	-1.2	2.5	0.1	-2.5	-4.0	-3.5	-5.5	0.4	-3.7
Y-displ (Å)	0.9	-0.4	1.3	0.4	-1.0	-2.4	-1.2	-0.6	-0.5	0.6
Inclination (°)	-18.8	-5.0	30.7	18.1	11.3	66.5	83.4	0.0	35.0	1.4
Tip (°)	-7.0	-8.7	2.5	-3.8	-9.1	22.7	45.4	-3.2	0.5	2.1
Global Inter-Base pair Parameters										
Shift (Å)	-0.0	0.1	-0.3	-0.1	0.0	0.8	1.4	0.2	-0.1	0.0
Slide (Å)	-0.3	0.3	0.5	0.1	0.5	-0.7	-0.6	0.3	0.1	0.3
Rise (Å)	3.8	3.8	5.5	4.1	3.5	2.9	2.7	3.1	4.8	3.3
Tilt (°)	-1.4	-1.5	-1.1	-2.5	1.6	4.1	3.3	2.3	0.6	1.2
Roll (°)	20.3	15.4	22.6	22.0	12.8	6.6	25.4	6.3	22.4	6.1
Twist (°)	16.4	-2.4	-17.2	0.2	-23.1	24.8	1.3	-23.9	-2.2	-28.1
Pucker	C1'exo C3'endo	North	C2'endo to C3'endo	C2'endo to C3'endo	C1'exo to C2'exo	C2'endo - C2'exo	Around C3'endo	Around C3'endo	Around C3'endo	Around C3'endo