Xylonucleic acid: synthesis, structure, and orthogonal pairing properties

Mohitosh Maiti1, Munmun Maiti1, Christine Knies2, Shrinivas Dumbre1, Eveline Lescrinier1, Helmut Rosemeyer2, Arnout Ceulemans3 and Piet Herdewijn1,*

1Laboratory of Medicinal Chemistry, Rega Institute for Medical Research, KU Leuven, Minderbroedersstraat 10, 3000 Leuven, Belgium, 2Organic Materials Chemistry and Bioorganic Chemistry, Institute of Chemistry of New Materials, University of Osnabrück, Barbarastrasse. 7, D-49069 Osnabrück, Germany and 3Department of Chemistry, KU Leuven, Celestijnenlaan 200F, 3001 Leuven, Belgium

ABSTRACT
There is a common interest for studying xeno-nucleic acid systems in the fields of synthetic biology and the origin of life. In particular, those with an engineered backbone and possessing novel properties. Along this line, we have investigated xylonucleic acid (XyloNA) containing a potentially prebiotic xylose sugar (a 3′-epimer of ribose) in its backbone. Herein, we report for the first time the synthesis of four XyloNA nucleotide building blocks and the assembly of XyloNA oligonucleotides containing all the natural nucleobases. A detailed investigation of pairing and structural properties of XyloNAs in comparison to DNA/RNA has been performed by thermal UV-melting, CD, and solution state NMR spectroscopic studies. XyloNA has been shown to be an orthogonal self-pairing system which adopts a slightly right-handed extended helical geometry. Our study on one hand, provides understanding for superior structure-function (pairing) properties of DNA/RNA over XyloNA for selection as an informational polymer in the prebiotic context, while on the other hand, finds potential of XyloNA as an orthogonal genetic system for application in synthetic biology.

INTRODUCTION
DNA and RNA are the two monophasic biopolymers that are capable of storage and dissemination of genetic information in the extant biology. While the nucleobases of DNA and RNA are at the very core of the helical structure providing crucial stabilization by Watson–Crick type hydrogen bonding and stacking interactions, the unique sugar-phosphate backbone is significantly providing an optimal structural and dynamic framework. These two natural nucleic acid versions differ in the sugar moiety in their backbone i.e. deoxy-D-ribose for DNA and D-ribose for RNA (Figure 1), thereby remarkably influencing their structural properties and biological functions (1). Recently, a number of sugar-modified nucleic acid variants has been revealed as new genetic polymers, (2) some of them are endowed with catalytic activity (for e.g. FANA and HNA) (3). The structure of these artificial nucleic acids, however, mimics natural nucleic acid helicity (4). In the context of studies on chemical etiology of nucleic acid structure (5–7) and the origin of life, alternative nucleic acids with more deviating backbone structures have been investigated. Some of them are having different double helical geometry than that of DNA/RNA (5–8) others with a similar geometry, and thus shown as potential direct RNA precursors (9–11) (like TNA and PNA) that might have been evolved into the existing nucleic acid. A relatively much less explored area in alternative nucleic acid engineering is the search and study of a complementary pair of synthetic nucleic acids with novel backbone and helical geometries which can support life i.e. perform catalysis like RNA and store genetic information like DNA. This might assist in the design principles of an orthogonal nucleic acid pair for application in xenobiology (orthogonal life) (12–14). Likewise, it may also expand our knowledge in the context of prebiotic origin and evolution of nucleic acids.

In this context, we have studied the physicochemical properties of a nucleic acid containing D-xylose (wood sugar), a prebiotic pentofuranosyl sugar. When compared with ribose, the xylose sugar has an inverted 3′-carbon center; hence xylose nucleic acid (XyloNA) and its cousin deoxyxylonucleic acid (dXyloNA) possess a diastereoisomeric backbone (Figure 1) in comparison with the natural RNA and DNA respectively. Previously, dXyloNA (15–18) and locked XyloNA (19–22) were studied in the perspective of potential therapeutical applications, but without success. Recently, we have shown the potential heterogeneous helical structure and conformational malleability of dXyloNA (23,24). In the present study, we describe the systematic ex-
were obtained by distillation over CaH₂. All other dry solvents and reagents were purchased from commercial sources. Pre-coated aluminum sheets (Fluka) were used for TLC. Flash silica column chromatography was performed on silica gel 60 A, 40–60 µm (Acros Organics). NMR spectra were recorded on Bruker Avance II spectrometers; 300 MHz with a 5 mm broad band probe, 500 MHz with a TXI-HCP Z gradient probe, or on a 600 MHz NMR equipped with a 5 mm TCI-HCN Z gradient cryo-probe. The spectra were processed with Bruker Topspin 2.1 software. Chemical shifts (δ) were reported in parts per million (ppm). The 1H and 13C chemical shifts were referenced relative to TMS peak (δ = 0.00 ppm). 31P NMR chemical shifts were referenced to an external 85% H3PO4 standard (δ = 0.00 ppm). 1H and 13C resonance assignments were made by using 2D correlation experiments (COSY, gHSQC and gHMBC). High resolution mass spectra were acquired on a quadrupole orthogonal acceleration time-of-flight mass spectrometer (Synapt G2 HDMS, Waters, Milford, MA, USA).

Synthetic procedures and the characterization data for the intermediates of XyloNA building blocks synthesis are provided in the supplementary information.

General procedure for the phosphitylation reaction
Trititated nucleoside (1 mmol) was coevaporated with dry THF (3×) and further dried under high vacuum for 1 h. Under an argon atmosphere, the compound was dissolved in anhydrous CH₂Cl₂ (10 ml) and N,N-diisopropyl ethylamine (5 mmol) was added. To this solution, 2-cyanoethyl-N,N-diisopropylchlorophosphoramidite (2.5 mmol) was added dropwise and the reaction mixture was stirred at room temperature for 4 h. Methanol (0.5 ml) was added dropwise to the reaction mixture at 0°C and the solution was stirred for another 15 min at room temperature. Then volatile organic solvents were removed under reduced pressure at room temperature and the crude product was purified by flash column chromatography.

1-{2'-O-Acetyl-3'-O-[2-cyanoethoxy(diisopropylami nophosphanyl)]-5'-O-(4,4'-dimethoxytrityl)-β-D-x ylofuranosyl}uracil (6)
Trititated nucleoside 5 (1.3 g, 2.2 mmol) was converted into the corresponding amidite by following the aforementioned general procedure for phosphitylation reaction. The crude product was purified by flash column chromatography using solvent system 25–75% EtOAc in hexane containing 0.5% triethylamine (R₁ = 0.49, 0.64 for 50% EtOAc in CH₂Cl₂). Compound 6 was isolated as white foam containing a mixture of two diastereomers in 87% yield (1.54 g). 1H NMR (500 MHz, acetone-d₆): δ 10.06 (br s, H₃), 7.63, 7.58 (d, J = 8.2 Hz, 1H, H6), 7.54–6.87 (m, 13H, Ar-H), 6.00, 5.99 (s, 1H, H1'), 5.54, 5.50 (d, J = 8.2 Hz, 1H, H5), 5.39, 5.23 (s, 1H, H2'), 4.53–4.49 (m, 1H, H4'), 4.43–4.40 (m, 1H, H3'), 3.94–3.70 (m, 2H, OCH₂CN), 3.80, 3.79 (s, 6H, 2 OCH₃), 3.69–3.31 (m, 2H, H3''), 3.54–3.43 (m, 2H, 2×CH(CH₃)₂), 2.73–2.56 (m, 2H, OCH₂CNCH₂), 2.12 (s, 3H, COCH₃), 1.15, 1.13, 1.02, 0.97 (d, T₂H, J = 5.0, 4.7, 6.7 Hz, 2×CH₂CH₃), 13C NMR (125 MHz, acetone-d₆): δ 169.9, 169.8 (COCH₃), 163.4, 163.3 (C4), 159.7 (Ar-C), 151.1 (C2), 146.0 (Ar-C), 141.0, 140.8 (C6), 136.9–127.6 (Ar-C), 118.8, 118.7 (CN), 114.0 (Ar-C), 102.3, 71.9acetone-d₆)

MATERIALS AND METHODS

General
For all reactions, analytical grade solvents were used. All moisture sensitive reactions were carried out under argon or nitrogen atmosphere in oven dried glasswares (120°C). 1,2-O-Isopropylidene-α-D-xylofuranose was purchased from Acros Organics. Dry acetonitrile and 1,2-dichloroethane were obtained by distillation over CaH₂. All other dry solvents and reagents were purchased from commercial sources. Pre-coated aluminum sheets (Fluka) were used for TLC. Flash silica column chromatography was performed on silica gel 60 A, 40–60 µm (Acros Organics). NMR spectra were recorded on Bruker Avance II spectrometers; 300 MHz with a 5 mm broad band probe, 500 MHz with a TXI-HCP Z gradient probe, or on a 600 MHz NMR equipped with a 5 mm TCI-HCN Z gradient cryo-probe. The spectra were processed with Bruker Topspin 2.1 software. Chemical shifts (δ) were reported in parts per million (ppm). The 1H and 13C chemical shifts were referenced relative to TMS peak (δ = 0.00 ppm). 31P NMR chemical shifts were referenced to an external 85% H₃PO₄ standard (δ = 0.00 ppm). 1H and 13C resonance assignments were made by using 2D correlation experiments (COSY, gHSQC and gHMBC). High resolution mass spectra were acquired on a quadrupole orthogonal acceleration time-of-flight mass spectrometer (Synapt G2 HDMS, Waters, Milford, MA, USA).

Synthetic procedures and the characterization data for the intermediates of XyloNA building blocks synthesis are provided in the supplementary information.

General procedure for the phosphitylation reaction
Trititated nucleoside (1 mmol) was coevaporated with dry THF (3×) and further dried under high vacuum for 1 h. Under an argon atmosphere, the compound was dissolved in anhydrous CH₂Cl₂ (10 ml) and N,N-diisopropyl ethylamine (5 mmol) was added. To this solution, 2-cyanoethyl-N,N-diisopropylchlorophosphoramidite (2.5 mmol) was added dropwise and the reaction mixture was stirred at room temperature for 4 h. Methanol (0.5 ml) was added dropwise to the reaction mixture at 0°C and the solution was stirred for another 15 min at room temperature. Then volatile organic solvents were removed under reduced pressure at room temperature and the crude product was purified by flash column chromatography.

1-{2'-O-Acetyl-3'-O-[2-cyanoethoxy(diisopropylami nophosphanyl)]-5'-O-(4,4'-dimethoxytrityl)-β-D-x ylofuranosyl}uracil (6)
Trititated nucleoside 5 (1.3 g, 2.2 mmol) was converted into the corresponding amidite by following the aforementioned general procedure for phosphitylation reaction. The crude product was purified by flash column chromatography using solvent system 25–75% EtOAc in hexane containing 0.5% triethylamine (R₁ = 0.49, 0.64 for 50% EtOAc in CH₂Cl₂). Compound 6 was isolated as white foam containing a mixture of two diastereomers in 87% yield (1.54 g). 1H NMR (500 MHz, acetone-d₆): δ 10.06 (br s, H₃), 7.63, 7.58 (d, J = 8.2 Hz, 1H, H6), 7.54–6.87 (m, 13H, Ar-H), 6.00, 5.99 (s, 1H, H1'), 5.54, 5.50 (d, J = 8.2 Hz, 1H, H5), 5.39, 5.23 (s, 1H, H2'), 4.53–4.49 (m, 1H, H4'), 4.43–4.40 (m, 1H, H3'), 3.94–3.70 (m, 2H, OCH₂CN), 3.80, 3.79 (s, 6H, 2 OCH₃), 3.69–3.31 (m, 2H, H3''), 3.54–3.43 (m, 2H, 2×CH(CH₃)₂), 2.73–2.56 (m, 2H, OCH₂CNCH₂), 2.12 (s, 3H, COCH₃), 1.15, 1.13, 1.02, 0.97 (d, T₂H, J = 5.0, 4.7, 6.7 Hz, 2×CH₂CH₃), 13C NMR (125 MHz, acetone-d₆): δ 169.9, 169.8 (COCH₃), 163.4, 163.3 (C4), 159.7 (Ar-C), 151.1 (C2), 146.0 (Ar-C), 141.0, 140.8 (C6), 136.9–127.6 (Ar-C), 118.8, 118.7 (CN), 114.0 (Ar-C), 102.3, 71.9acetone-d₆)
1-(2′-O-Acetyl-3′-O-[2-cyanoethoxy(diisopropylamino)phosphoryl]5′-O-(4,4′-dimethoxytrityl)-β-d-xylofuranosyl)-4-N-benzoylcytosine (10)

Tritiated nucleoside 9 (1.35 g, 2.0 mmol) was converted into the corresponding amide by following the general procedure for phosphorylation. The crude product was purified by flash column chromatography using solvent system 20–40% acetone in hexane containing 0.5% triethylamine (Rt = 0.68, 0.66 for 60% acetone in hexane). Compound 10 was isolated as white foams containing a mixture of two diastereoisomers in 81% yield (1.4 g). 1H NMR (500 MHz, acetone-d6): δ 89.78 (br s, NH), 8.16, 8.15 (d, J = 6.9, 6.5 Hz, 2H, o-H-Bz), 7.97, 7.96 (d, J = 7.4 Hz, 1H, H6), 7.68–7.64 (m, 1H, p-H-Bz), 7.58–7.56 (m, 2H, m-H-Bz), 7.44–7.33 (d, J = 7.4 Hz, 1H, H5), 7.58–6.90 (m, 13H, Ar-H of DMTr), 5.94, 5.91 (s, 1H, H1′), 5.53, 5.54 (s, 1H, H2′), 4.72–4.65 (m, 1H, H4′), 4.39–4.37 (m, 1H, H3′), 3.82, 3.81 (s, 6H, 2OCH3), 3.82–3.54 (m, 2H, OCH2CH2CN), 3.82–3.33 (m, 2H, H5′, β′), 3.40–3.33 (m, 2H, 2′xCH2CH2CN), 2.71–2.56 (m, 2H, OCH2CH2CN), 2.15 (s, 3H, COCH3), 1.09, 1.04, 0.92 (d, 12H, J = 6.8 Hz, 2′xCH3(CH2)2); 13C NMR (125 MHz, acetone-d6): δ 169.8, 169.7 (COCH3), 164.0, 163.9 (C4), 159.8 (Ar-C), 155.2 (C2), 146.1 (Ar-C), 145.2 (C6), 136.8–127.7 (Ar-C), 118.8, 118.7 (CN), 114.0, 113.9 (Ar-C), 96.3 (C5), 92.1, 91.9 (C1′), 87.3, 87.2 (tort C of DMTr), 84.9, 84.8 (d, Jcp = 4.5 Hz, C4′), 80.8, 80.2 (d, Jcp = 3.3, 0.0 Hz, C2′), 77.0, 75.8 (d, Jcp = 19.3, 11.6 Hz, C3′), 63.9, 63.7 (C5′), 59.9, 59.5 (d, Jcp = 23.1, 22.9 Hz, OCH2CH2CN), 55.6 (OCH3), 43.9, 43.8 (d, Jcp = 9.4, 8.8 Hz, CH2(CH2)2), 25.1, 24.9, 24.6 (d, Jcp = 8.2, 7.3, 4.9 Hz, CH2(CH2)2), 20.8, 20.7 (COCH3), 20.6, 20.2 (d, Jcp = 7.5, 8.2 Hz, OCH2CH2CN); 31P NMR (202 MHz, acetone-d6): δ 152.80, 148.52; HRMS (ESI+), calculated for C46H34N4O10P [M+Na]+ 938.3613, found 938.3616.

9-(2′-O-Acetyl-3′-O-[2-cyanoethoxy(diisopropylamino)phosphoryl]5′-O-(4,4′-dimethoxytrityl)-β-d-xylofuranosyl)-6-N-benzoyladenine (14)

Tritiated nucleoside 13 (2.28 g, 3.2 mmol) was converted into the corresponding amide by following the general procedure for phosphorylation. The crude product was purified by flash column chromatography using solvent system 45–85% EtOAc in hexane containing 0.5% triethylamine (Rt = 0.42, 0.54 for 90% EtOAc in hexane). Compound 14 was isolated as white foams containing a mixture of two diastereoisomers in 84% yield (2.46 g). 1H NMR (500 MHz, acetone-d6): δ 10.00 (br s, NH), 8.67, 8.65 (s, 1H, H2), 8.30, 8.29 (s, 1H, H8), 8.14, 8.12 (d, J = 6.2, 6.9 Hz, 2H, o-H-Bz), 7.66–7.63 (m, 1H, p-H-Bz), 7.57–7.53 (m, 2H, m-H-Bz), 7.57–6.85 (m, 13H, Ar-H of DMTr), 6.40, 6.39 (s, 1H, H1′), 5.82, 5.68 (s, 1H, H2′), 4.70–4.69 (m, 1H, H4′), 4.59–4.53 (m, 1H, H3′), 3.79, 3.78, 3.77 (s, 6H, 2OCH3), 3.74–3.54 (m, 2H, OCH2CH2CN), 3.74–3.35 (m, 2H, H5′, β′), 3.42–3.35 (m, 2H, 2′xCH2CH2CN), 2.65–2.52 (m, 2H, OCH2CH2CN), 2.18, 2.17 (s, 3H, COCH3), 1.11, 1.08, 0.95, 0.92 (d, 12H, J = 6.9, 7.2, 6.9, 6.7 Hz, 2′xCH2CH2CN); 13C NMR (125 MHz, acetone-d6): δ 169.9 (COCH3), 159.7 (Ar-C), 152.8 (C4′), 151.2 (C6), 146.1 (Ar-C), 142.5, 142.4 (C8), 136.9–127.6 (Ar-C), 125.9, 125.7 (C5), 118.8, 118.7 (CN), 113.9 (Ar-C), 89.1, 88.7 (C1′), 87.3, 87.2 (tort C of DMTr), 84.0, 83.9 (d, Jcp = 3.4, 0.0 Hz, C4′), 81.3, 81.2 (d, Jcp = 2.5, 0.0 Hz, C2′), 77.3, 76.0 (d, Jcp = 18.1, 12.5 Hz, C3′), 64.2, 64.0 (C5′), 59.8, 59.5 (d, Jcp = 21.9, 20.8 Hz, OCH2CH2CN), 55.6 (OCH3), 43.9, 43.8 (d, Jcp = 12.5, 12.1 Hz, CH(CH2)2), 25.1, 24.9, 24.7, 24.6 (d, Jcp = 8.0, 7.5, 8.2, 8.0 Hz, CH2(CH2)2), 20.8, 20.7 (COCH3), 20.5, 20.4 (d, Jcp = 7.5 Hz, OCH2CH2CN); 31P NMR (202 MHz, acetone-d6): δ 152.14, 148.98; HRMS (ESI+), calculated for C46H34N4O10PNa [M+Na]+ 938.3613, found 938.3616.
spectroscopic studies are provided in the supplementary information.

NMR-derived structural restraints

Distance restraints were derived from NOESY spectra recorded with 50, 100, 150 and 300 ms mixing times. By using the computer-aided resonance assignment (CARA) program, (25) the cross-peaks in NOESY spectra were assigned and the cross-peak volumes were obtained by integration. Based on the NOE build-up curves, inter-proton distances were calculated. The calibration of NOE cross-relaxation rates was performed against the average of all the isolated H5-H6 cross-peaks as a reference with distance relaxation rates was performed against the average of all distances were calculated. The calibration of NOE cross-relaxation. Based on the NOE build-up curves, inter-proton distances were calculated. The calibration of NOE cross-relaxation rates was performed against the average of all the isolated H5-H6 cross-peaks as a reference with distance relaxation rates was performed against the average of all distances were calculated. The calibration of NOE cross-relaxation.

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

Backbone torsion angles β, γ and ε were restrained as follows. The β torsion angles (C4′-C5′-O5′-P5′) were restrained to trans regions in the structure calculations as determined by experimental vicinal hetero-nuclear phosphorus proton (3JPS-H5S, 3JPS-H4S) couplings (Supplementary Table S1, in the supporting information) derived from the 31P coupled and decoupled DQF-COSY spectrum and based on a standard Karplus plot (26). For purine and pyrimidine nucleotides of XyloNA1 duplex, the β torsion angles were restrained to 20° ± 20° and 20° ± 40° respectively. The γ torsion angles (C3′-C4′-C5′-O5′) were restrained from gauche – to anticalinal – region in the structure calculations as determined by vicinal proton-proton (3JH4-H5S, 3JH4-H4S) couplings (Supplementary Table S1) derived from the 31P decoupled DQF-COSY spectrum (26). For purine and pyrimidine nucleotides of XyloNA1 duplex, the γ torsion angles were restrained to 210° ± 35° and 270° ± 45° respectively. The ε torsion angles (P3′-O3′-C3′-C4′) were restrained to the trans region (ε = 180°±60°) in all the residues of XyloNA1 based on the observed cross-peaks for long range phosphorus proton couplings (4JH4-P4) in 1H-31P HETCOR spectra (Supplementary Figure S1, in the supporting information), indicative for a W-shaped conformation of the P3′-O3′-C3′-C4′-H4′ fragment in the backbone.

No torsion angle restraints were applied on the backbone dihedral angles α, and ζ.

Structure calculations

All structure calculations were performed using the X-PLOR NIH program (27). The standard topology and parameter file (topallhdg.dna, parallhdg.dna) for DNA were adapted for Xylose 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 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 during a ‘gentle molecular dynamics’ round by using the X-PLOR NIH script refine.gentle.inp. The final refinement was started with 20 ps constant-temperature molecular dynamics simulation at 300 K (20 000 steps of 0.001 ps) which gradually introduces the van der Waals radii, 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 (28). Finally, visual representations of the molecules were obtained by using the UCSF Chimera 1.10rc program.

RESULTS

Synthesis of XyloNA building blocks

In the chemistry part, we have set out to implement an easy synthetic route (Scheme 1) for the XyloNA building blocks. The commercially available 1,2-acetone of xylofuranose was benzylated to have compound 1. The subsequent steps to obtain protected glycone (2) are acid-catalyzed acetonide deprotection and acetylation. Since the one-pot transformation solely produced a ring-opened byproduct 2b, we have followed a two-step procedure to synthesize diacetate 2 in a reasonably good yield. Although the formation of the 1-fluoroacetate side product 2a could not be completely avoided in the latter procedure, its formation was reduced to some extent by decreasing the percentage of trifluoroacetic acid (TFA). Fluoroacetate sugar 2a could not afford the desired nucleoside 3 in a classical glycosylation reaction since it formed several decomposition products as observed by TLC. While standard Vorbrüggen-type glycosylation reaction was performed for pyrimidine nucleosides (3 and 7), for purines non-standard conditions were necessary. Thus, the reaction of N-benzyladenine with the protected sugar 2 was performed in the presence of SnCl4 without silylation at room temperature, where the desired N′-β-isomer (11) was majorly formed as a kinetically controlled product. Nonetheless, depending on the reaction time (<2 h) a minor amount of N′-α-isomer (11a in SI) was formed and separated from the N′-β-stereoisomer 11 by column chromatography. For guanine, N′-acetyl-O′-diphenylcarbamoyl (DPC) protection (15 in SI) was employed for the regioselective glycosylation at the N′ position, (29) since glycosylation without DPC protection (15a in SI) resulted in the formation of an inseparable (1:1) mixture of N′ (17a in SI) and N (17) regioisomers. Then the DPC protecting group was selectively removed by treating with 90% aqueous TFA as it was no longer necessary and thought to procure complication in the subsequent steps. The removal of the benzyl (-Bn) groups was found to be difficult due to the sluggish reaction rate under different hydrogenation conditions, with potential problems like reduction of the pyrimidine ring and solvolysis of the acetyl and benzoyl groups were additionally encountered. These problems were efficiently
addressed by using the Lewis-acid (BCl3) mediated debenzylation reaction. The debenzylated xylonucleosides were tritylated and phosphitylated to the corresponding amidites (Table 1). Three fully modified oligonucleotide sequences (Table 1, XyloNA1–3) containing β-D-xylofuranosyl nucleotides were assembled on the solid support using synthesized phosphoramidites.

Thermal (UV-) melting study

The fully modified self-complementary XyloNA oligonucleotide (ON-1, Figure 1) was initially investigated for its capability in forming a duplex structure by thermal (UV-) melting study. This data indicates that XyloNA is able to form a self-paired duplex with a melting temperature (Tm) of 67°C and is associated with a typical sigmoidal cooperative melting profile (Figure 2A). In order to compare the thermal stability of duplexes possessing different backbones with the same sequence context (Table 1), we have synthesized and analyzed ON-1 to ON-4. The relative thermal stability of different duplexes decreased in the order (Table 1, ON1–4) of XyloNA: XyloNA:dXyloNA; RNA:RNA; DNA:DNA. Notably, the xylose-based backbone modification (XyloNA/dXyloNA) significantly increased the duplex stability by 1.6–2.2-fold in comparison with the corresponding ribose-based DNA/RNA pair. In particular, the effect of the 2′-hydroxyl moiety is instrumental in increasing the duplex stability of XyloNA versus dXyloNA, as also seen in the case of natural RNA vs. DNA system. A non self-complementary fully-modified XyloNA oligonucleotide with mixed sequence context (Table 1, XyloNA2) was additionally synthesized and investigated for its potential to cross-pair with the natural DNA/RNA complements. These studies highlight that the XyloNA system could neither pair with DNA nor with RNA (Table 1 and Figure 2B, C), and thus demonstrating autonomous self-pairing property of XyloNA.

In order to investigate if XyloNA is universally orthogonal or not in pairing with other DNA/RNA sequence, we have also constructed a homo-pyrimidine sequence (XyloNA3, uU12) and evaluated its cross pairing ability with natural DNA (dA12) and RNA (rA12) complements. As shown in Figure 3, XU12 is indeed capable of pairing with both dA12 and rA12 strands (Table 1). In both the cases triplex formation could be confirmed based on the following observations (Figure 3): (i) a strong hysteresis effect (ΔTm = 13°C between heating and cooling curves) was observed (Figure 3) in normal melting buffer, in particular with rA12, which can be considered unusual for a duplex; (ii) these structures are highly stabilized in a Mg2+ buffer without any significant hysteresis effect; since the extent of cationic stabilization is usually more pronounced in a triplex than a duplex; (iii) moreover, a temperature-dependent hypochromism was observed in 284 and 295 nm (Figure 3) which represents the melting of Hoogsteen paired third-strand from a triplex structure (30–32). Melting temperatures of both the strands of these triplexes are identical and represent a single transition at 240 nm while melting of the third strand can only be visible at 284/295 nm. Melting...
Circular dichroism (CD) study

To gain insights into the helical conformation and dynamics of double-stranded (ds) XyloNA (ON-1), we have performed circular dichroism (CD) studies. The CD spectra of XyloNA appear to be different from the known CD spectra of various helical nucleic acid structures i.e. A-, B- and Z-type helices, as depicted in Figure 4. However, the CD spectrum of XyloNA was found to be very similar to that of our previously studied analogous dsXyloNA system in water (24). A characteristic positive Cotton (effect) band at 256 nm was observed in the CD spectrum, which showed a gradual melting effect with increasing temperature (Figure 4, bottom). These observations suggest that the XyloNA duplex has formed a slightly right-handed extended ladder-like structure, as characterized by NMR (described in the following section), and is in line with our previous observation for dsXyloNA (24). The repeated appearance of this type of CD spectrum can be considered as a fingerprint for the right-handed extended duplex structure (RE-type helix), and might facilitate in the future cataloging of non-natural nucleic acid conformation by CD spectroscopy. Unlike dsXyloNA duplex (24), the CD spectra of XyloNA in saline buffer (Supplementary Figure S5) appeared to be identical with the spectra in pure water. This suggests that the presence of the 2'-hydroxyl group in the XyloNA backbone has resulted in a reduced conformational flexibility compared to the dsXyloNA system. In our previous study, the latter has been identified to be conformationally malleable since it exists in two different helical conformations in saline aqueous/buffer solution (24).

Structure elucidation by solution state NMR spectroscopy

To examine the nature of the helical structure adopted by the dsXyloNA (ON-1), we have undertaken solution state NMR structural analysis. Appearance of a single set of resonances in the NMR spectra suggests the formation of a unique type of duplex conformation (Figure 5A and B). The duplex forms Watson–Crick type base-pairing with anti-parallel strand orientation as evidenced by the hydrogen-bonded imino-proton signals in the NMR spectrum recorded in H2O (Figure 5A and Supplementary Figure S6). All the xylose sugar rings are in N-type puckering as suggested by the appearance of singlet anomic-proton peaks in the 1H NMR spectrum (Figure 5B). Unlike native DNA/RNA double helices, unique diagonal cross-strand (n+1 to n) NOE interactions between anomic (H1') protons were observed in XyloNA self-duplex, as depicted in Figure 5C. Although less pronounced, a classical NOE walk (Figure 6) between aromatic and anomic protons was observed, that is generally more prominent in the natural DNA/RNA helices. During simulations, the structures were converged (structure statistics in SI, Supplementary Table S2) to a family of structures (Figure 7B-D and E) with similar geometries and energies. The aforementioned diagonal NOE distance restraints were also reflected in the final structure as depicted in Figure 7A. Diagnostic helical pa-
moidal behavior was observed and thus XyloNA2:DNA hybrid (ON-6), up- and down-curves are shown; no signification in XyloNA. High and dissimilar inclination together with largely reduced helicity of XyloNA compared with DNA/RNA might be the reasons for its inability in cross-pairing with the DNA/RNA complements (37). A comparison of the torsion angles (Table 2) between XyloNA and DNA/RNA suggests a cooperative adjustment in different dihedral angles across the sugar-phosphate backbone. The chiral inversion in the 3′-carbon center of the sugar ring in XyloNA with respect to DNA/RNA primarily changed the \textit{endo} cyclic torsion angle \( \delta \) to \textit{synclinal}° from \textit{gauche}° or \textit{anticlinal}°. This leads to a chain of inversions in \( \alpha, \gamma \) and \( \zeta \) torsion angles from an almost mirror-image domain (opposite in sign and magnitude, shown in Table 2) while retaining both the torsion angles \( \beta \) and \( \epsilon \) in the \textit{anti} domain. These distinctive features allow a rationalization of the structural orthogonality of XyloNA in comparison with the natural DNA/RNA.

### DISCUSSION

As outlined above, our endeavor in synthesizing and characterizing a xeno-nucleic acid analogue by appending the potentially prebiotic xylose sugar in the backbone has resulted in the creation of a novel xylonucleic acid (XyloNA) system with interesting structural and biophysical properties. Like RNA, XyloNA could have been formed and existed in the prebiotic era since xylose has shown to be a product in the formose reaction (38), the most plausible prebiotic route of sugar synthesis. This suggests a possibility that XyloNA/dXyloNA could have been selected by natural evolution as a genetic information system. There are, however, several structure-function properties of the xylonucleic-based information system (as revealed by our studies) that make it less suitable for this purpose. Although XyloNA/dXyloNA could form base-paired duplexes, it lacks the very helical nature and thus the compactness possessed by the natural helices. Moreover, XyloNA/dXyloNA systems, by virtue of their predominant interstrand stacking property, are thermodynamically more stable than the natural ones. The higher duplex stability might be a plausible disadvantage since it would be rather difficult to unwind the helices for accessing the genetic information in the prebiotic and/or biotic polymerization. Previously it was emphasized that natural DNA/RNA structures are selected based on the principle of optimization rather than maximization of duplex stability and conformational flexibility (5). The highly inclined bases along with reduced helicality in XyloNA/dXyloNA strand make them ineligible to cross-pair with the natural nucleic acids in the general sequence context, as evidenced by the melting studies (Table 1). Since unable to cross-pair, it lacks communication with the RNA world, and therefore XyloNA could not be considered as a direct prebiotic RNA progenitor (39). These unfavorable properties together might have contributed and favored the selection of RNA rather than XyloNA as an informational polymer in the prebiotic nucleic acid evolution. Besides structure-pairing properties, there is a prebiotic generational issue for any proposed primordial RNA analogues. Conformational rigidity and altered helical geometry of XyloNA in comparison with RNA may potentially influence the prebiotic polymerization fidelity, linkage regio-chemistry (3′-5′ versus 2′-5′ isomers) and differential
Figure 3. Thermal (UV-) melting profile for (A) pairing of XyloNA3 (xU12) with dA12; (B) pairing of XyloNA3 (xU12) with rA12. Sigmoidal melting curves are shown in different wavelengths (260, 284 and 295 nm). Heating- and cooling-curves are highlighted in red and blue respectively. Melting experiments were generally performed in 20 mM KH2PO4 buffer (pH 7.5) containing 0.1 M NaCl, and 0.1 mM Na2EDTA or in a Mg2+-buffer, as indicated above, containing 10 mM Tris·HCl (pH 7.5) and 50 mM MgCl2. Melting curves at 295 nm are from a faster heating-cooling experiments. Data presented here are only for a 1:1 mixing stoichiometry with 4 μM concentration of individual strand.

Table 2. Torsion angles and key helical parameters of natural B-DNA and A-RNA duplexes are in comparison with the synthetic XyloNA and dXyloNA duplex structures derived by NMR structural analysisa

<table>
<thead>
<tr>
<th>Torsion angles [°]</th>
<th>B-DNA</th>
<th>A-RNA</th>
<th>dXyloNA</th>
<th>XyloNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>– 47</td>
<td>– 62</td>
<td>53 ± 18</td>
<td>47 ± 20</td>
</tr>
<tr>
<td>β</td>
<td>– 146</td>
<td>180</td>
<td>– 156 ± 14</td>
<td>– 130 ± 14</td>
</tr>
<tr>
<td>γ</td>
<td>36</td>
<td>47</td>
<td>– 122 ± 8 (A/G)</td>
<td>– 100 ± 25 (A/G)</td>
</tr>
<tr>
<td>δ</td>
<td>156</td>
<td>83</td>
<td>– 64 ± 4 (C/T)</td>
<td>– 91 ± 9 (C/U)</td>
</tr>
<tr>
<td>ε</td>
<td>155</td>
<td>– 152</td>
<td>25 ± 3</td>
<td>21 ± 6</td>
</tr>
<tr>
<td>ζ</td>
<td>– 95</td>
<td>– 73</td>
<td>146 ± 13</td>
<td>132 ± 16</td>
</tr>
<tr>
<td>ξ</td>
<td>– 98</td>
<td>– 166</td>
<td>75 ± 12</td>
<td>86 ± 21</td>
</tr>
</tbody>
</table>

Helical parameters

| Twist [°] | 36.0 | 32.7 | 2.7 ± 0.5 | 10.7 ± 1.7 |
| Inclination [°] | – 5.9 | 15.8 | – 52.2 ± 0.4 | – 45.2 ± 0.9 |
| Inclination-τb [°] | 0 | – 30 | 51 | 41 |
| Rise [Å]   | 3.4  | 2.8  | 5.8 ± 0.06 | 5.1 ± 0.05 |

aStandard values for DNA, RNA and parameters for dXyloNA are from (24). Average and standard deviations (±) were calculated for the XyloNA duplex from the output data generated by using the CURVES program for an ensemble of 20 refined NMR structures. Backbone-base inclination angles (τb) were calculated by using the ‘Inclination’ program, (37).
chemical stability (40) of the competing regioisomers that can form, as shown previously in the case of polymerization of different natural and non-natural systems (41–44). In the prebiotic synthesis of nucleotide building blocks, there can also be a different outcome with a xylose nucleoside containing 2',3'-trans diol functionality than a ribonucleoside with 2',3'-cis diol functionality (22,45). These aspects remain to be studied to realize the full prebiotic potential of XyloNA.

Previously, six-membered sugar-modified homo-DNA/pyranosyl-RNA (p-RNA) (8) and acyclic glycol nucleic acid (GNA) (46) have also adopted a similar extended ladder-like structure. Interestingly, we could here demonstrate that an extended helical structure could also be derived by introducing a pentofuranosyl (Xylo) sugar. Xylose nucleic acid (XyloNA) is found to be conformationally less flexible (as seen in RNA compared to DNA), due to the 2'-hydroxyl effect, than the previously studied analogous dXyloNA counterpart (24). In contrary, the structural compactness and flexibility of XyloNA and dXyloNA are not as favorable as in the case of natural (DNA/RNA) systems. DNA (usually B-form) and RNA (A-form) can shape into different types of helical con-

Figure 4. Top: Circular dichroism (CD) spectra of XyloNA and dXyloNA duplexes in water (12 μM, pH 7.0) at 10°C. For comparison, the CD spectra of the right-handed B-DNA, A-RNA and the left-handed Z-DNA duplexes are shown and illustrated. Bottom: Temperature dependent CD spectra of the XyloNA duplex showing a typical cooperative melting profile.

Figure 5. (A) Hydrogen-bonded imino-proton peaks of the XyloNA duplex (ON-1) representing Watson–Crick type base-pairing; (B) Appearance of eight singlet anomic-proton peaks (G1 to C8) in the 1H NMR spectrum of the XyloNA duplex showing N-type puckering for the xylose sugar in XyloNA; (C) a unique diagonal cross-strand NOE interaction(s) between H1'-protons observed (boxed and indicated by red arrow) in the 2D NOESY spectrum of XyloNA duplex. All the above spectra were recorded in a 600 MHz spectrometer with a 1.5 mM XyloNA (ON-1) sample (pH 7.2) at 10°C.

Figure 6. Expansion of the NOESY spectrum (recorded with 300 ms mixing time, in a 600 MHz spectrometer) of XyloNA (ON-1: 5'-G'UGUACAC8–3') duplex (1.5 mM, pH 7.2) in D2O at 10°C, showing NOE walk between H6/H8-H1' intra-residue and H1'(n)-H6/H8(n+1) inter-residue protons.
formations by adopting different sugar pucker (1) (as shown in Figure 1), as well as by essentially changing the important endo-cyclic $\delta$-torsion angle in the backbone. A similar structural polymorphism could not be observed in the xylose-based nucleic acids due to an identical $\delta$ value and sugar pucker for both XyloNA and dXyloNA, and thereby limiting their conformational diversity. Torsion angle $\delta$ is found here to be crucial in determining the helical twist and backbone flexibility versus rigidity as previously suggested for DNA/RNA (47). Chiral inversion in the C3$'$ of furanose sugar in XyloNA/dXyloNA is constraining the sugar conformation and $\delta$ torsion angle in comparison with DNA/RNA, as a result forming a ladder-like structure with increased duplex stability and to some extent rigidifying the backbone conformation (16). Although dXyloNA was previously shown to be structurally somewhat malleable in its duplex conformation than natural DNA duplex (23,24). In DNA/RNA, duplex stability and backbone flexibility seem to have optimally selected and supported by their very unique double helical structures. Considering the favorable properties that are required for an efficient storage and transfer of genetic information, natural DNA/RNA structure is unique and superior to that of xylose-based (XyloNA/dXyloNA) nucleic acids.

Although xylonucleic acid could not supersede the structure and pairing properties of natural nucleic acids, its backbone could support essential self-pairing ability to form double helix with slightly right-handed extended conformation. It is anticipated that XyloNA should be chemically and enzymatically more stable than RNA due to an inverted C-3$'$ orientation of the phosphate linkage, and thus making the phosphodiester moiety inaccessible by the neighboring 2$'$-hydroxyl nucleophile. By virtue of its unique structure, it could demonstrate novel orthogonal (inability in pairing with DNA/RNA) self-pairing properties. Triplex formation with purine DNA/RNA (as shown in melting study) is a special case with pyrimidine XyloNA/dXyloNA sequences. From an applicability point of view, if XyloNA/dXyloNA is not a universal orthogonal system, most importantly XyloNA/dXyloNA (as shown for XyloNA2) should be an orthogonal information system in a mixed sequence context which is mostly the case with natural unbiased genomic sequences. In general, xylose or deoxyxylose nucleic acid by virtue of their inability to cross talk with the natural nucleic acids could in principle be used for construction of an orthogonal episome for exploration in synthetic biology (12–14). Along this line, the main bottleneck for implementing such a XyloNA-based technology is the discovery of a XyloNA polymerase which can efficiently polymerize XyloNA nucleotide precursors. In our previous study, dXylo nucleoside triphosphates were shown to be partially recognized by some mutant DNA polymerases, like for example Taq mutant M1, for up to four nucleotides incorporation (24). Recently, polymerases have been evolved that can copy or transfer the genetic information from the non-natural nucleic acid backbones, those mimicking the natural helical structure (2,3). Polymerase evolution for XyloNA like de-viant nucleic acid system poses a major challenge in the hands of evolutionary molecular biologist.

Figure 7. Structural representations of the XyloNA duplex as determined by NMR structural analysis, (A) a close view of the minimum energy closest-to-average structure showing base stacking interactions and the proximity of the diagonal anomeric protons as observed in the NOESY spectra, (B) view from the minor groove, (C) side view with highlighted backbone, (D) view from the major groove, and (E) an overlay of 20 refined structures representing convergence in the structure calculation.
CONCLUSION
Studies on artificial nucleic acid analogues from the structural neighborhood of DNA/RNA are inspired by motivations like gaining knowledge on the structural uniqueness of DNA/RNA, and understanding the primordial ancestors that are lost in the evolutionary past. Among them, those with novel properties could be useful in synthetic biology. Herein, we have investigated such an isomeric RNA analogue, XyloNA derived from prebiotic xylose sugar. A convergent synthetic route was implemented to access XyloNA building blocks. Structural and pairing properties of the synthesized XyloNAs are elucidated by employing biophysical techniques. The helical form of DNA/RNA is found to be structurally unique in a sense that they possess optimal duplex stability and conformational flexibility in comparison with that of XyloNA. Unlike DNA/RNA, XyloNA adopts an altered extended helical geometry, and thus unable to cross-pair with the unbiased sequences of DNA/RNA. Based on the preliminary structure-function understanding available so far, it seems that RNA may have gained an evolutionary advantage over XyloNA. On the other hand, the autonomous pairing potential of XyloNA could further be exploited as an orthogonal information system in synthetic biology.

SUPPLEMENTARY DATA
Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS
Dr Mohitoshi Maiti is a postdoctoral fellow (1200113N) of FWO Vlaanderen, Belgium. Technical support by Mr Guy Schepers is gratefully acknowledged. We thank Dr Puntee Srivastava for scientific discussion and Chantal Biernaux for secretarial assistance. This article is dedicated to Professor A. Eschenmoser on the occasion of his 90th birthday.

FUNDING
European Research Council under the European Union’s Seventh Framework Programme [FP7/2007–2013] /ERC Grant agreement [ERC-2012-ADG_20120216/320683]; FWO Vlaanderen, Belgium (G078014N to P.H., O6260 to MoM); Mass spectrometry was made possible by the support of the Hercules Foundation of the Flemish Government [20100225-7]. Funding for open access charge: FWO Vlaanderen, Belgium [O6260 of MoM].
Conflict of interest statement. None declared.

REFERENCES