JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

Phosphonomethyl Oligonucleotides as Backbone Modified Artifi-cial Genetic Polymers

Chao Liu, Christopher Cozens, Faten Jaziri, Jef Rozenski, Amandine Maréchal, Shrinivas G. Dumbre, Valérie Pezo, philippe marliere, Vitor B. Pinheiro, Elisabetta Groaz, and Piet Herdewijn

J. Am. Chem. Soc., **Just Accepted Manuscript** • DOI: 10.1021/jacs.8b03447 • Publication Date (Web): 03 May 2018

ACS Publications

is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036

Phosphonomethyl Oligonucleotides as Backbone Modified Artificial Genetic Polymers

Chao Liu,[†,⊥] Christopher Cozens,[§,⊥] Faten Jaziri,[‡] Jef Rozenski,[†] Amandine Marechal,[§] Shrinivas Dumbre,[†] Valérie Pezo,[‡] Philippe Marlière,[‡] Vitor B. Pinheiro,[§,∥,*] Elisabetta Groaz,[†] and Piet Herdewijn[†,‡,*]

[†]Medicinal Chemistry, Rega Institute for Medical Research, KU Leuven, Herestraat 49, 3000 Leuven, Belgium

[§]University College London, Gower Street, London WC1E 6BT, UK

[‡]*i*SSB, Genopole, CNRS, UEVE, Université Paris-Saclay, 5 rue Henri Desbruères, 91030 Evry Cedex, France

[∥]Institute of Structural and Molecular Biology, Birkbeck College, University of London, Malet Street, WC1E 7HX, UK

[⊥]These authors contributed equally to this work.

KEYWORDS: XNA; Phosphono nucleic acid; Polymerase evolution; In vivo transliteration; H-Phosphinates; Xenobiology

ABSTRACT: Although several synthetic or xenobiotic nucleic acids (XNAs) have been shown to be viable genetic materials *in vitro*, major hurdles remain for their *in vivo* applications, particularly orthogonality. The availability of XNAs that do not interact with natural nucleic acids and are not affected by natural DNA processing enzymes, as well as specialized XNA processing enzymes that do not interact with natural nucleic acids, are essential. Here, we report 3'–2' phosphonomethyl-threosyl nucleic acid (tPhoNA) as a novel XNA genetic material and a prime candidate for *in vivo* XNA applications. We established routes for the chemical synthesis of phosphonate nucleic acids and phosphorylated monomeric building blocks, and we demonstrated that DNA duplexes were destabilized upon replacement with tPhoNA. We engineered a novel tPhoNA synthetase enzyme and, with a previously reported XNA reverse transcriptase, demonstrated that tPhoNA is a viable genetic material (with an aggregate error rate of approximately 17 x 10^{-3} per base) compatible with the isolation of functional XNAs. *In vivo* experiments to test tPhoNA orthogonality showed that the *E. coli* cellular machinery had only very limited potential to access genetic information in tPhoNA. Our work is the first report of a synthetic genetic material modified in both sugar and phosphate backbone moieties and represents a significant advance in biorthogonality towards the introduction of XNA systems *in vivo*.

INTRODUCTION

In RNA and DNA, life on Earth has settled on uniquely efficient molecules to store and, facilitated by highly efficient enzymes, propagate genetic information. While other molecules (XNAs, xenobiotic nucleic acids)[1] can fulfil the function of those natural genetic polymers, establishing them as bona fide genetic materials is challenging. It requires the efficient chemical synthesis of unnatural precursors as well as isolation of polymerases that can, at least, transfer genetic information in and out of a natural system efficiently.[2,3] It is especially important to acquire efficient chemical methods to synthesize primers for polymerase selection as well as specific oligonucleotide sequences in order to establish whether XNA polymers could support duplex formation with a complementary XNA, DNA, or RNA strand and exhibit canonical Watson-Crick pairing. While XNAs are established as synthetic genetic polymers in vitro,[4] only very few studies have been conducted to evaluate backbone modified XNAs *in vivo*.[5-9] So far, the replication of such XNA systems has been accomplished through a DNA intermediate, which is not compatible with the *in vivo* generation of an orthogonal XNA episome. Notably, the study of backbone modified XNAs remains underexplored compared to that of base modified XNAs.[10-13] Moreover, a limited metabolic stability, lack of orthogonality, and toxicity further restrict *in vivo* applications of XNAs, highlighting the need for alternative designs in this area.

Among all the XNA genetic systems studied to date, (3′,2′)-α-L-threose nucleic acid (TNA, Figure 1) has generated considerable interest as a possible RNA progenitor or competitor because of the chemical simplicity of threose relative to ribose and the ability of TNA to form stable duplexes with RNA and DNA, despite having a repeating backbone unit one atom shorter than that of natural nucleic acids.[14-16] Storage of genetic information in TNA has been established through synthesis of TNA oligomers by Therminator DNA polymerase[17] and reverse transcription by several natural and engineered enzymes.[4, 18, 19]

Previously, our group demonstrated that Therminator polymerase was also able to catalyze the condensation of the diphosphate derivatives of both phosphonomethylthreosyl-adenine (PMTApp) and 5'-*O*-phosphonomethyl-2'-deoxyadenine (PMdApp) with the formation of phosphonate oligonucleotide stretches, which were found to be resistant to degradation by snake-venom phosphodiesterase.[20,21] The solid-phase synthesis of 5'-*O*-phosphonomethyl-deoxyribosyl oligonucleotides (dPhoNA, Figure 1) was first reported by Rosenberg *et al.* via the phosphotriester method.[22] The as-synthesized phosphonate oligonucleotides exhibited high

stability against nucleases of L1210 cell free extract and possessed the ability to form triplexes with natural counterparts. In a recent study, oligothymidylates containing various ratios of 5'-*O*-methylphosphonate internucleotide linkages were found to be able to induce RNA cleavage by ribonuclease H (RNase H).[23]

The main objective of this study was to investigate the potential of 3'–2' phosphonomethyl-threosyl nucleic acid (tPhoNA) as novel XNA genetic material for *in vivo* applications, particularly in comparison with dPhoNA. The presumed enzymatic and chemical stability of the phosphonate linkage, which would avoid degradation by cellular enzymes, along with a significant advance in biorthogonality due to the concurrent modification of both sugar and phosphate backbone moieties are significant advantages of those chemistries. First, synthetic routes were established for the preparation of phosphonate oligonucleotides as well as their corresponding phosphorylated building blocks required for chemical and enzymatic synthesis. We then characterized the influence of tPhoNA and dPhoNA on duplex stability, establish tPhoNA as a synthetic genetic polymer through polymerase engineering, permitting enzymatic tPhoNA synthesis and reverse transcription, and finally demonstrated tPhoNA's orthogonality by *in vivo* transliteration experiments.

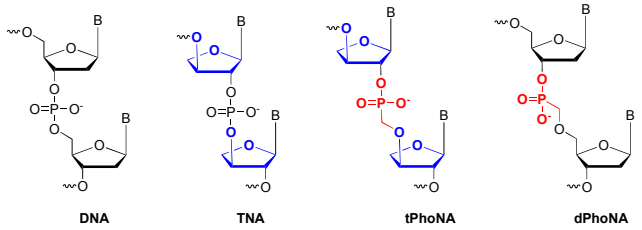


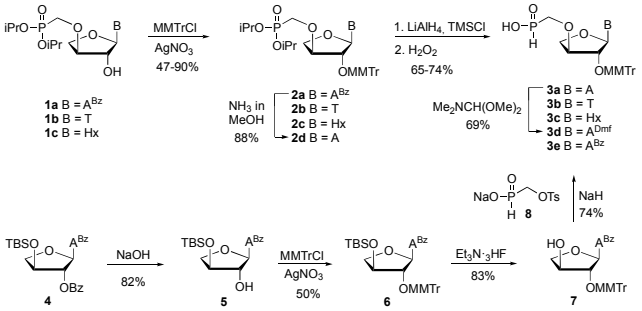
Figure 1. Chemical structures of phosphono nucleic acids (tPhoNA and dPhoNA) as compared to DNA- and TNA-based phosphodiester backbones.

RESULTS AND DISCUSSION

Phosphonate Nucleic Acid Synthesis. *Chemical Synthesis of Nucleoside H-Phosphinate Monomers.* As depicted in Scheme 1, threosyl *H*-phosphinate monomers **3b-d** were prepared starting from 3'-diisopropylphosphonomethyl threosyl nucleosides **1a** (B = benzoyl adenine, A^{Bz}), **1b** (B = thymine, T), and **1c** (B = hypoxanthine, Hx), respectively, which were in turn synthesized following literature procedures starting from L-ascorbic acid.[24,25] Compounds **1a-c** were reacted with MMTrCl in the presence of $AgNO_3$ and pyridine to afford the corresponding 2'-*O*-tritylated nucleoside analogues **2a-c** in 47–90% yield. In view of the incompatibility of the benzoyl protecting group with the conditions required by the following $LiAlH_4$ mediated reduction step, compound **2a** was first deprotected at the N^6-position by treatment with methanolic ammonia to give adenine derivative **2d**. Next, the diisopropyl ester phosphonate functionality of **2b-d** was reduced to the corresponding phosphine stage in the presence of $LiAlH_4$-TMSCl, followed by oxidation with hydrogen peroxide to afford threosyl *H*-phosphinate nucleosides **3a-c** in 65–70% yield over two steps. The free exocyclic amino group of adenine-containing nucleoside **3a** was then reprotected with a dimethylformamidino (dmf) group to provide *H*-phosphinate monomer **3d**.

Alternatively, the adenine *H*-phosphinate building block containing a benzoyl protected adenine base, compound **3e**, could be obtained by directly introducing the *H*-phosphinate functionality into the sugar moiety of threose nucleoside intermediate **7**, as described by Kostov et al. (Scheme 1).[26] Selective hydrolysis of the benzoyl group at the 2'-position of fully protected nucleoside **4**[24] under basic conditions gave compound **5** in 82% yield. Protection of the 2'-hydroxyl group of **5** with MMTrCl afforded nucleoside **6** in 50% yield, along with the recovery of 42% of the starting material. The subsequent removal of the 3'-*tert*-butyldimethylsilyl protecting group was carried out by treatment of **6** with triethylamine trifluoride in THF. Next, compound **7** was reacted with sodium 4-toluene-sulfonyloxymethyl-(*H*)-phosphinate **8**[26] in the presence of sodium hydride in DMF to give *H*-phosphinate nucleoside **3e** in 74% yield. Finally, all these *H*-phosphinate monomers were transformed to their DBU salts prior to solid-phase synthesis. In addition, for comparison with phosphonate-linkage modified oligonucleotides including a natural sugar in the backbone, 5'-*O*-phosphonomethyl-deoxyribosyl oligonucleotides (dPhoNA) were also prepared and likewise 2'-Deoxyribonucleoside-5'-*O*-methyl-(*H*)-phosphinates (**S4b-e**) were synthesized starting from deoxyribonucleosides in a similar manner as that described for threosyl *H*-phosphinate monomers (Scheme S1).

Scheme 1. Synthetic routes for the preparation of threosyl nucleoside H-phosphinate monomers 3b-e required for the solid-phase synthesis of tPhoNA containing oligonucleotides. Similar synthetic steps were employed for the preparation of deoxyribosyl nucleoside H-phosphinate building blocks (S4b-e) for dPhoNA synthesis, as detailed in the SI (Scheme S1).



Synthesis of tPhoNA-DNA and dPhoNA-DNA Chimeras. Recently, Rosenberg et al. reported the preparation of a fully modified 7-mer dPhoNA oligomer via *H*-phosphinate chemistry employing 2-chloro-5,5'-dimethyl-1,3,2-dioxaphosphorinane-2-oxide (NEP-Cl) and an anhydrous mixture (CCl_4–methanol–Et_3N–MeIm) as coupling and oxidizing agents, respectively.[26] Here, we opted for the synthesis of PhoNA oligonucleotides and their DNA chimeras using a combination of *H*-phosphinate and conventional phosphoramidite chemistries. However, different conditions were adopted along the synthesis cycle compared to those previously reported, as detailed in the Supporting information (Scheme S3, Table S1).

Specifically, considering that the *H*-phosphinate linkage is sensitive to conventional water-containing oxidizers (I_2 in pyridine/H_2O), it was immediately oxidized to a methyl group protected phosphonate linkage after the condensation step.

Second, the oxidation step was followed by a capping step using phenoxyacetic anhydride to avoid failure sequences. Third, 33% aqueous ammonia was preferred as final deprotection reagent over previously reported conditions (gaseous NH_3).[26]

As proof of principle, a fully modified polythymidylate 14-mer dPhoNA oligomer could be obtained by using this modified synthesis protocol. Stability studies on this oligomer revealed that the phosphonate linkage (P-C-O) was stable against aqueous alkaline deprotection conditions (Figure S1). Next, up to six successive threose phosphonate nucleotides were successfully incorporated into DNA oligomers by this method. In addition, two 10-mer sequences each flanked by one DNA unit at both ends were likewise obtained (Table 1, ON**9** and ON**10**). The mass analysis and sequences of the synthesized oligomers are listed in Table 1. Oligonucleotides ON **1**, ON **3**, ON **11**, and ON **14** were prepared using A^{Dmf} protected nucleotides, while ON **5**–**10** and ON **13** were attained using A^{Bz} protected nucleotides.

Table 1. Mass Analysis of PhoNA-DNA Chimeras Used for T_m Studies.[a]

ON	Sequence	Calc.	Obs.
ON **1**	5'-TTT TAA **A**TA TAA-3'	3640.7	3640.6
ON **2**	5'-TTA TA**T** TTA AAA-3'	3640.7	3640.9
ON **3**	5'P-TGC ATG GC**A** CGG CGC TAG-3'	5617.9	5618.2
ON **4**	5'P-CTA GCG CCG **T**GC CAT GCA-3'	5537.9	5538.1
ON **5**	5'-CTA GCG CCG **TAT** CAT GCA-3'	5457.0	5457.0
ON **6**	5'-TGC ATG **ATA** CGG CGC TAG-3'	5537.0	5537.0
ON **7**	5'-CTA GCG **ATT TAT** CAT GCA-3'	5471.0	5471.0
ON **8**	5'-TGC ATG **ATA AAT** CGC TAG-3'	5520.0	5520.0
ON **9**	5'-T**TA TAT TTA AA**A-3'	3640.7	3640.7
ON **10**	5'-T**TT TAA ATA TA**A-3'	3640.7	3640.7
ON **11**	5'-TTT TAA **a**TA TAA-3'	3654.7	3654.8
ON **12**	5'-TTA TA**t** TTA AAA-3'	3654.7	3654.8
ON **13**	5'-T**tt taa ata taa**-3'	3794.8	3794.9
ON **14**	5'-T**ta tat tta aaa**-3'	3794.8	3794.9

[a]Modified units are bolded and underlined. tPhoNA and dPhoNA nucleotides are shown as blue capital and magenta lowercase letters, respectively.

Chemical Synthesis of Diphosphates of Threosyl Nucleoside Phosphonates (PMTNpp). The four PMT diphosphates with A, T, C (cytosine), and G (guanine) bases required for polymerase engineering (Figure 2) were obtained from the corresponding nucleoside phosphonic acids (**S5a-d**, Scheme S2), which were synthesized according to literature reports.[24,27] Previously, the diphosphate derivative of phosphonomethyl-threosyl adenine (PMTApp) was isolated in poor yield (10%) by using the *N,N'*-carbonyldiimidazole (CDI) approach.[20] Therefore, we decided to prepare all PMTNpp using an alternative dicyclohexylcarbodiimide (DCC) mediated condensation method,[28] which pleasingly led to an improvement of the diphosphorylated phosphonate product yields (30–40%).

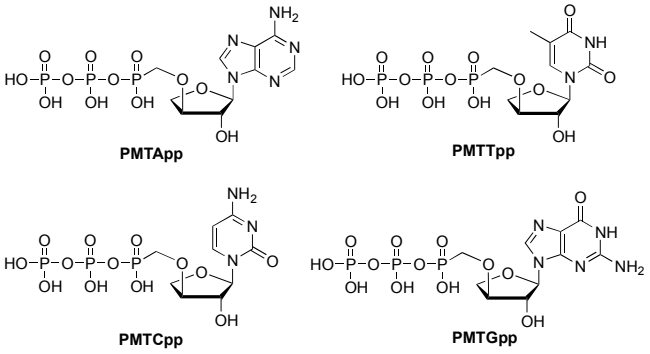


Figure 2. Chemical structures of diphosphates of threosyl nucleoside phosphonates (PMTNpp) serving as substrates for engineering of a dedicated tPhoNA polymerase through Tgo directed evolution.

Hybridization Properties of tPhoNA and dPhoNA. *Heteroduplex Stability with DNA.* In order to evaluate the hybridization properties of phosphonate-linked modified oligonucleotides, the thermal stabilities of duplexes composed of PhoNA-DNA chimeras (ONs **1–14**) and complementary strands were evaluated by temperature-dependent UV spectroscopy. The T_m values of the corresponding unmodified duplexes were determined for comparison.

Generally, the introduction of a 3'-*O*-phosphonomethyl-threosyl unit into a DNA sequence resulted in a considerable destabilization of the DNA/DNA duplex relative to the corresponding unmodified duplex with ΔT_m values ranging from -1.6 to -7.0 °C (Table 1, ONs **1–4**). As the number of modifications increased from one to three and six (ONs **5–8**), the T_m dropped within a range from -0.6 to -5.2 °C, while showing a modest reduction in ΔT_m per modification (-0.2 to -1.3 °C/mod). When 10 modifications were introduced to produce a T/A 12 mer (ON **9** and **10**), the T_m decreased below the measurable threshold. In comparison, the introduction of a 5'-*O*-phosphonomethyl-deoxyribosyl unit containing an adenine base (ON **11**) resulted in a thermal stability similar to that of the unmodified counterpart (ΔT_m = +0.1 °C/mod.), although the insertion of one 5'-*O*-phosphonomethyl-deoxyribosyl unit bearing a thymine base (ON **12**) resulted in a considerable destabilization of the DNA/DNA duplex (ΔT_m = -2.4 °C/mod.). The fully modified 5'-phosphonate 12-mer oligomers (ON **13** and ON **14**) showed no hybridization with a complementary DNA strand.

Table 2. T_m Data for Duplexes of tPhoNA and dPhoNA Containing Oligonucleotides with their DNA Complements.

tPhoNA				
Duplex with DNA	N° of modifications	T_m (°C)[a]	ΔT_m (°C)[b]	ΔT_m/mod (°C)
ON **1**	1	31.2	- 1.6	- 1.6
ON **2**	1	26.9	- 5.9	- 5.9
ON **3**	1	66.5	- 7.0	- 7.0
ON **4**	1	71.2	- 2.3	- 2.3
ON **5**	3	64.2	- 3.8	- 1.3
ON **6**	3	67.4	- 0.6	- 0.2
ON **7**	6	55.0	- 5.2	- 0.9
ON **8**	6	56.0	- 2.0	- 0.3
ON **9**	10	< 5	> - 27.8	> - 2.8
ON **10**	10	< 5	> - 27.8	> - 2.8
dPhoNA				
Duplex with DNA	N° of modifications	T_m (°C)[a]	ΔT_m (°C)[b]	ΔT_m/mod (°C)
ON **11**	1	32.9	+ 0.1	+ 0.1
ON **12**	1	30.4	- 2.4	- 2.4
ON **13**	11	< 5	> - 27.8	> - 2.5
ON **14**	11	< 5	> - 27.8	> - 2.5

dsDNA Reference Duplexes and T_m			
Sequence	T_m (°C)	Sequence	T_m (°C)
5'-TTTTAAATATAA-3' 3'-AAAATTTATATT-5'	32.8	5'-CTA GCG CCG TAT CAT GCA-3' 3'-GAT CGC GGC ATA GTA CGT-5'	68.0
5'p-TGC ATG GCA CGG-CGC TAG-3' 3'-ACG TAC CGT GCC GCG ATC-p5'	73.5	5'-CTA GCG ATT TAT CAT GCA-3' 3'-GAT CGC TAA ATA GTA CGT-5'	60.2

[a] T_m measurements represent an average of at least two experiments. [b] ΔT_m versus dsDNA duplexes. [c] Rejman et al.[22] described ΔT_m/mod between -0.2 °C (T^*_{15}:dA_{15}) and -2.4 °C (A^*_{15}:dT_{15}) for homopolymers with a triple helix motif.

Heteroduplex Stability with RNA. Hybridization studies on tPhoNA-DNA chimeras with RNA demonstrated that the ON **5**/RNA duplex (T_m = 65.1 °C, Table 3) was more stable relative to its ON **5**/DNA duplex (64.2 °C). As observed for the ON **5**/RNA duplex, the ΔT_m was -0.8 °C for three modifications introduced when compared to the corresponding DNA/RNA duplex. Increasing the number of modifications from three to six (ON **7**, Table 3) resulted in a moderate decline in T_m from 54.8 to 53.7 °C with a ΔT_m value of -0.2 °C per modification. On the contrary, the fully modified dPhoNA oligomer (ON **14**, Table 3) could form a duplex with a complementary strand of RNA (T_m = 29.0 °C) that was more stable as compared to the DNA/RNA duplex (T_m = 19.6 °C) and slightly less stable compared to the RNA/RNA duplex (T_m = 31.7 °C). Remarkably, when the sequence was inverted between both strands (ON **13**, Table 3), the duplex with RNA became less stable.

Homoduplex Formation of tPhoNA and dPhoNA. In addition, in order to establish whether phosphonate oligonucleotides could form self-complementary duplexes, the hybridization properties of PhoNA-DNA chimeras with complementary PhoNA-DNA strands were studied. The results showed that the introduction of one pair of phosphonate nucleotides (ON **1**/ON **2** and ON **3**/ON **4**, Table 4) led to a considerable destabilization of the DNA/DNA duplex (ΔT_m = -7.2 °C and -7.5 °C, respectively). However, extension of the modifications from one to three and six pairs (ON **5**/ON **6** and ON **7**/ON **8)** did not lead to a steep decline in T_m. Instead, a similar reduction in T_m was observed with ΔT_m values of -1.1 and -1.5 °C per modification, respectively. On the contrary, the introduction of one pair of dPhoNA nucleotides (ON **11**/ON **12**) resulted in a moderate destabilization of the DNA/DNA duplex (ΔT_m = -0.9 °C), while the fully modified dPhoNA oligomer ON **14** could not form a duplex with its complementary strand. The 10-mer tPhoNA (ON **9**/ON **10)** displayed a broad melting curve, suggesting that the melting process is not an all-or-none melting profile. The association between both complementary oligonucleotides (ON **9**/ON **10**) is of undefined nature.

The above T_m data obtained for dPhoNA and dPhoNA-DNA chimeras with complementary strands of DNA and RNA is generally in accordance with previous data reported by Rosenberg et al.[22,23] However, in both cases (tPhoNA-DNA and dPhoNA-DNA chimeras), a higher binding affinity for a complementary RNA stand as opposed to a DNA strand was observed. This could arise from the different conformations of the resultant PhoNA-DNA/DNA and PhoNA-DNA/RNA duplexes.[29,30] Based on experiments with T/A mixed sequences, it is not clear whether tPhoNA could form stable homodu-

plexes, although the properties of this association was not analyzed.

Table 3. T_m Data for Duplexes of tPhoNA and dPhoNA Containing Oligonucleotides with their RNA Complements.

tPhoNA				
Duplex with RNA	N° of modifications	T_m (°C)[a]	ΔT_m (°C)	ΔT_m/mod (°C)
ON **5**	3	65.1	- 0.8 (b)	- 0.3
ON **7**	6	53.7	-1.1 (b)	- 0.2
ON **9**	10	< 5	> - 26.7 (c)	> - 2.7
ON **10**	10	< 5	> - 26.7 (c)	> - 2.7
dPhoNA				
Duplex with RNA	N° of modifications	T_m (°C)[a]	ΔT_m (°C)[b]	ΔT_m/mod (°C)
ON **13**	11	12.7	- 6.4 (b)	- 0.6
ON **14**	11	29.0	+ 9.4 (b)	+ 0.85

DNA/RNA and dsRNA Reference Duplexes and T_m (°C)

Sequence	T_m (°C)	Sequence	T_m (°C)
5'-d(TTTTAAATATAA)-5' 3'-r(AAAAUUUAUAUU)-5'	16.2	5'-d(CTA GCG CCG TAT CAT GCA)-3' 3'-r(GAU CGC GGC AUA GUA CGU)-5'	65.9
5'-d(TTA TAT TTA AAA)-3' 3'-r(AAU AUA AAU UUU)-5'	19.6	5'-d(CTA GCG ATT TAT CAT GCA)-3' 3'-r(GAU CGC UAA AUA GUA CGU)-5'	54.8
5'-r(UUU UAA AUA UAA)-3' 3'-r(AAA AUU UAU AUU)-5'	31.7		

[a] T_m measurements represent an average of at least two experiments. [b] ΔT_m versus DNA/RNA duplexes, (c) ΔT_m versus RNA/RNA duplexes.

Table 4. T_m Data for Duplexes of tPhoNA and dPhoNA Containing Oligonucleotides: Self-Hybridization of Chimeric Duplexes Containing DNA.

tPhoNA				
	N° of modifications	T_m (°C)[a]	ΔT_m (°C)[b]	ΔT_m/mod (°C)
ON **1**/ON **2**	1	25.6	- 7.2	- 7.2
ON **3**/ON **4**	1	66.0	- 7.5	- 7.5
ON **5**/ON **6**	3	64.6	- 3.4	- 1.1
ON **7**/ON **8**	6	51.4	- 8.8	- 1.5
ON **9**/ON **10**	10	31.7 (c)	- 1.1	- 0.1
dPhoNA				
	N° of modifications	T_m (°C)[a]	ΔT_m (°C)[b]	ΔT_m/mod (°C)
ON **11**/ON **12**	1	31.9	- 0.9	- 0.9
ON **13**/ON **14**	11	< 5	> - 27.8	> - 2.5

[a] T_m measurements represent an average of at least two experiments. [b] ΔT_m versus dsDNA. [c] broad melting profile.

Establishing tPhoNA as a Xenobiotic Genetic Material. Many *in vitro* and *in vivo* applications can only be realized if XNA oligomers are proven to be viable genetic materials, capable of storing and propagating genetic information accessible by XNA polymerases.[31] As PMTNpp were poor substrates for commercial DNA polymerases (Figure S2), and in view of the significant basal incorporation activity for PMTApp exhibited by Therminator[20] (an A485L mutant of the archaeal 9°N DNA polymerase), we decided to engineer a tPhoNA synthase from the same starting enzyme previously used for the directed evolution of HNA synthases (TgoT, a Therminator homologue).[4] TgoT is a Tgo (*Thermococcus gorgonarius*) DNA polymerase variant harboring four mutations: V93Q,[32] D141A, E143A, and A485L.[33] Since archaeal B-family polymerases have been extensively engineered for XNA synthesis and reverse transcription,[34] we chose a stepwise approach to polymerase engineering, testing previously reported mutations and screening residues known to be implicated in XNA substrate specificity (summarized in Figure S3).

First, we introduced mutations known to improve sugar-modified XNA synthesis (L403P)[4] and enhance polymerase processivity (H147E).[35] L403 is adjacent to the highly conserved polymerase A motif (D404 – P410) and the L403P

mutation has been shown to improve incorporation of arabino (and fluoroarabino) nucleic acids.[4] The H147E mutation affects the probability of nascent strand switching between polymerase and exonuclease active sites, reducing exonuclease and increasing primer extension activities. Both L403P (TgoT_P) and the double mutant (TgoT_EP: H147E, L403P) displayed improved tPhoNA synthesis compared to TgoT, and were capable of full-length synthesis (57 PMTpp incorporations) from a mixed-sequence template (TempN, Table S6) under forcing conditions (high polymerase concentration, long extensions, supplemented with manganese), as shown in Figure S3.

We then sought to improve polymerase processivity by optimizing nascent nucleic acid binding. Mutations to residue E664 on the polymerase thumb have been shown to be essential for efficient RNA (E664K)[36] and TNA (E664I)[3] synthesis as well as improved 2'-5' backbone modified DNA synthesis.[37] Screening all possible 20 E664 variants for tPhoNA synthesis identified E664H (TgoT_EPH) as a clear improvement over TgoT_EP (Figure 3 and S5).

Next, we introduced I521L, a mutation originally isolated for XNA reverse transcription[4] but shown to enhance XNA synthesis when added to some XNA synthetases.[38] The resulting mutant (TgoT_EPLH) showed modest gains in activity for both PMTNpp (Figure S5) and dNTP incorporation (Figure S4).

Finally, comparing recent KOD DNA polymerase ternary complex structures with natural (dATP)[39] and unnatural threosyl (tATP)[40] substrates suggested that Tyr409 was not well-oriented to provide stabilizing stacking interactions with the threosyl sugar of the incoming tATP nucleotide. We speculated that larger hydrophobic moieties at L408 could better position Y409 for stacking against threosyl sugars, stabilizing PMTNpp binding and improving its incorporation. Both mutations tested (L408I and L408F) increased tPhoNA synthetase activity, especially L408F. The resulting polymerase (TgoT_EPFLH: H147E, L403P, L408F, I521L, E664H) was found to be an efficient tPhoNA synthetase (Figure 3A), capable of synthesizing a 57-mer (from TempN template) in less than five minutes (Figure S7) using low polymerase (0.2 μM) and manganese (100 μM) concentrations. In fact, full-length tPhoNA synthesis could be achieved even in the absence of manganese (Figure S5). The synthesized polymer showed higher mobility in denaturing gels (Figure S7) and increased exonuclease resistance (Figure S8) compared to DNA. Enzymatic synthesis of tPhoNA was confirmed by far infrared spectroscopy, where an observed frequency change of the (PO_2-) ν_{as} from 1225-1240 to 1205 cm^{-1} appeared as one of several consistent markers for the bond change from P-O-C in DNA to P-C-O in tPhoNA (Figure 3D and Figure S10).

TgoT_EPFLH retained some of the biases observed in the parental polymerase, such as poor PMTTpp homopolymeric incorporation (Figure S7A). In mixed sequences, two sequential PMTTpp incorporations reproducibly led to polymerase stalling (Figure S7B) and were avoided from subsequent experiments. The aggregate fidelity of PMTTpp incorporation did not significantly differ from PMTApp incorporation, which showed no stalling pattern, suggesting that the stalling is due to problems accommodating consecutive PMTTpp incorporations, rather than catastrophic loss of fidelity (Figure S9A and S9C). An unintended consequence of this systematic engineering has been that every mutation (apart from I521L) introduced to improve tPhoNA synthesis resulted in reduced DNA synthesis capability (Figure S4). It remains to be shown if this shift in substrate is the result of a specialist DNA polymerase being converted into a generalist XNA polymerase, or if Tgo_EPFLH engineering is demonstrating the first steps towards the engineering of orthogonality.

Having synthesized tPhoNA, we screened available reverse transcriptases (RTs) that could synthesize DNA against XNA templates. The previously described T.RT521 (TgoT: I521L)[4] and a close homologue harboring equivalent mutations K.RT-521K (KOD DNA polymerase: V93E, D141A, E143A, A485L, I521L, E664K) were both efficient tPhoNA RTs capable of DNA synthesis from both DNA and RNA primers (Figure 3B and Figure S6), despite the poor hybridization observed between tPhoNA and DNA in the absence of polymerase, akin to the previously observed reverse transcription of glycerol nucleic acids.[41]

Together, Tgo_EPFLH and both RTs tested enabled us to transfer genetic information from DNA into tPhoNA and recover that information back to DNA (Figure 3C), with an approximate aggregate error rate of 17-20 x 10^{-3} per incorporation (K.RT521K and T.RT521, respectively – Figure S9); a degree of fidelity compatible with the development of aptamers and aptazymes based on this chemistry.[4,38]

Analysis of the errors (Figure S9), showed a significant bias towards the misincorporation of PMTGpp (or misincorporation of dCTP during RT). A similar effect has been observed in TNA synthesis with T.RT521 and could be minimized through the use of biased unnatural nucleotide pools lower in the guanosine analogue.[4]

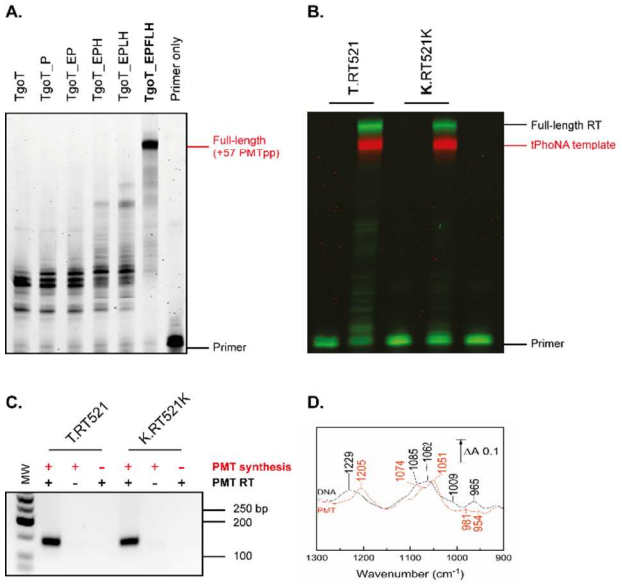


Figure 3. tPhoNA as a genetic polymer. (A) Impact of the systematic engineering of the starting polymerase (TgoT) for the synthesis of tPhoNA under forcing conditions (Figure S5 for details). Intermediates in the engineering process including TgoT_P (TgoT: L403P), TgoT_EP (TgoT_P: H147E), Tgo_EPH (TgoT_EP: E664H), TgoT_EPLH (TgoT_EPH: I521L) as well as the isolated tPhoNA synthase TgoT_EPFLH (TgoEPH: L408F) are shown. (B) Reverse transcription of tPhoNA to DNA. Both T.RT521 and K.RT521K were capable of template-dependent DNA synthesis (green) against a chimer-

ic tPhoNA template (20 nt DNA from the original synthesis primer with 39 PMT incorporations based on TempN2), which is shown in red, from an RNA primer. (C) PCR amplification of the tPhoNA RT products was only possible when both tPhoNA synthesis and reverse transcription were carried out, confirming that tPhoNA can store and propagate genetic information. (MW: NEB Low molecular weight DNA ladder). (D) ATR-FTIR spectra of DNA (black) and tPhoNA (red) in the 1300-900 cm^{-1} range. Frequency changes of the (PO_2^-) ν_{as} from 1225-1240 to 1205 cm^{-1}, a shift of the (PO_2^-) symmetric stretching mode (ν_s) at 1085 cm^{-1}, and the (C-O) ribose stretch at 1062 cm^{-1} in DNA to 1074 and 1051 cm^{-1} in tPhoNA can be observed as well as signal reductions at 1009 cm^{-1} (assigned to a ribose ring mode) and the characteristic ribose-phosphodiester skeletal motion mode of DNA at 965 cm^{-1}. Complete spectra are shown in Figure S10.

***In Vivo* PhoNA-Dependent DNA Synthesis.** To evaluate the ability of phosphonate modified nucleic acids to serve as template for DNA synthesis in *E. coli,* six 5'-phosphorylated DNA-PhoNA-DNA chimera oligonucleotides (ON **4** and **15-19**, Table S3) were synthesized and tested using the established gapped-vector assay based on the restoration of the active site of thymidylate synthase (Figure 4).[5,42] This enzyme catalyzes the conversion of deoxyuridine monophosphate to thymidine monophosphate, and is essential for *E. coli* growth in a medium lacking either thymine or thymidine (dt). The six 18-base long 5'-phosphorylated DNA-PhoNA-DNA chimera oligonucleotides were ligated into a gapped heteroduplex plasmid where 6 codons surrounding the catalytic Cys146 of the *thyA* gene had been deleted. The resulting ligation products were then transformed into an *E. coli* strain lacking *thyA*. Transformants are only able to survive in thymidineless media when a PhoNA oligonucleotide chimera is recognized by the bacterial replication machinery and utilized as template for DNA synthesis, restoring the thymidylate synthase active site. The ratio between bacterial colony numbers in media with or without thymidine indicates the extent of successful templating.

As expected, ligation with no oligonucleotide or of a Cys146-deleted oligonucleotide did not yield any prototrophic transformants. The replacement of a DNA unit by either a tPhoNA or dPhoNA building block (ON **4** and **17**, respectively) yielded a similar 2.5-fold decrease of the number of prototrophic transformants compared to the positive controls (Tables S4 and S5, Figure 4B and 4C, respectively). A further 2- and 6.5-fold drop in the yield of prototrophic transformants resulted from extension of the tPhoNA stretch from one to two (ON **15**) and three (ON **16**) oligonucleotides, respectively. Thus, the successive addition of tPhoNA nucleotides significantly diminished DNA propagation *in vivo*. Even more appreciable yield drops were observed upon extension of the 2'-deoxyribose-5'-phosphonate stretch from one to two oligonucleotides (ON **18**, 76-fold decrease) and from one to three oligonucleotides (ON **19**, 217-fold decrease). These results suggest that the dPhoNA backbone modification caused less productive *in vivo* transliteration than tPhoNA by *E. coli* DNA polymerases. This is presumably due to the difference in length of the internucleotide linkage between tPhoNA (6 bonds per repeating unit, in analogy to DNA) and dPhoNA (7 bonds per repeating unit), as depicted in Figure 1. Since replacement of guanine in the Cys146 codon (TGC/T) by an hypoxanthine is well tolerated and does not interfere with the correct copying of DNA,[5] we interpret the decrease of active *thyA* genes among the clones transformed upon extending the phosphonate nucleotide stretch as evidence of inefficient copying by *E. coli* polymerases.

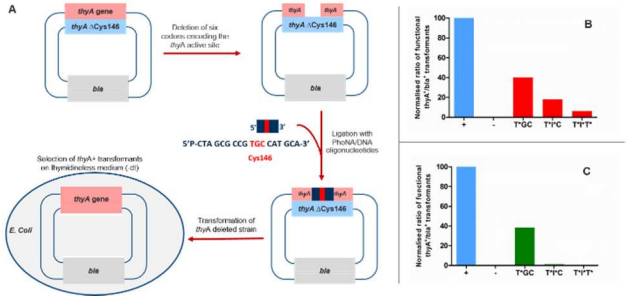


Figure 4. (A) Selection screening of PhoNA containing oligonucleotides capable of templating DNA synthesis in *E. coli* based on the restoration of the *thyA* gene.[5,42] *In vivo* propagation of DNA-tPhoNA-DNA (B) and DNA-dPhoNA-DNA (C) chimeras. T* and I* indicate Pho-modified units with a thymine and hypoxanthine base, respectively. The normalized ratios correspond to the experimentally derived average number of thymidine prototrophic colonies ($thyA^+$) from the average total number of colonies (bla^+).

CONCLUSION

For *in vivo* applications, it is desirable that XNAs are both chemically and biologically orthogonal: that is that neither the polymers nor the building blocks interact with natural nucleic acids or proteins, and that XNA-synthesizing and binding proteins do not synthesize or bind natural nucleic acids. Significant progress has been made in recent years in describing XNA oligomers and establishing them as genetic polymers through engineered polymerases that can synthesize and reverse transcribe to transfer genetic information from DNA to XNA and back. However, most of these XNAs retain significant affinity for DNA and RNA and substantial efficacy in *in vivo* transliteration experiments. Furthermore, all engineered XNA polymerases so far described also retain substantial DNA polymerase capability.

Thus, we set out to evaluate for the first time chemically modified nucleic acids featuring a phosphonate internucleotide linkage for their potential as orthogonal genetic materials. Specifically, 3'–2' phosphonomethyl-threosyl (tPhoNA) oligonucleotides, along with 5'–3' phosphonomethyl–deoxyribosyl (dPhoNA) sequences as well as their DNA chimeras were conveniently prepared by *H*-phosphinate and phosphoramidate solid-phase syntheses upon careful adaptation of the synthesis cycle conditions.

Both chemistries showed significant levels of orthogonality at both chemical (oligonucleotide properties) and biological (recognition by natural protein) levels. Our melting analyses demonstrated that heavily modified tPhoNA and dPhoNA (ON **9**/**10** and **13**/**14**, respectively) did not exhibit detectable hybridization to complementary DNA (and, for tPhoNA, RNA) strands. Significantly, tPhoNA retained some potential to form homoduplexes (ON **9**/**10**, Table 4), at least for AT-rich sequences.

tPhoNA also showed signs of biological orthogonality. PMTNpps were demonstrably poor substrates for natural polymerases (Figure S1), yet tPhoNA could be efficiently

synthesized by the engineered polymerase TgoT_EPFLH (Figure 3). Importantly, as we engineered TgoT for better tPhoNA synthesis, we observed a noticeable drop in its DNA synthetase function (Figures S3 and S4). Given the orthogonality demonstrated both chemically (oligonucleotide annealing) and *in vivo* (transliteration), it is likely that TgoT_EPFLH's broadened substrate specificity can be further engineered to develop an orthogonal polymerase.

Together, these data suggest that a fully orthogonal genetic system based on tPhoNA and specialist tPhoNA polymerases, and with minimal interaction with natural dNTPs, nucleic acids or polymerases, is achievable.

METHODS

General Chemistry Details. Detailed procedures for the synthesis of threosyl and deoxyribosyl nucleoside *H*-phosphinates as well as diphosphates of threosyl nucleoside phosphonates are described in detail in the Supporting Information. All reactions were carried out under an argon atmosphere using oven-dried glassware. Chemical reagents were purchased from commercial sources and used as received. ^{1}H, ^{13}C, and ^{31}P spectra were recorded on a 300, 500, or 600 MHz Bruker Avance spectrometer. 2D NMR experiments (H-COSY, HSQC, and HMBC) were used for the characterization of all the intermediates and final compounds. High resolution mass spectra were recorded on a quadrupole orthogonal acceleration time-of-flight mass spectrometer (Synapt G2 HDMS, Waters, Milford, MA). Preparative HPLC purification was performed on a Phenomenex Gemini 110A column (C18, 10 μm, 21.2 mm × 250 mm). Column chromatography was performed using silica gel 60 Å, 0.035−0.070 mm (Acros Organics).

Solid-Phase Synthesis. The solid-phase synthesis of PhoNA oligonucleotides as well as PhoNA-DNA chimeras was accomplished with an Expedite DNA synthesizer (Applied Biosystems) by using either an adapted *H*-phosphinate approach or its combination with the standard phosphoramidite method (Scheme S3, Table S1). After synthesis, the support was dried in vacuo, treated with a freshly prepared thiophenol-Et_3N-DMF mixture (23:32:45; v/v) for 4 h to remove the methyl protecting groups at the phosphonate linkages, rinsed with dry acetonitrile, and further dried in vacuo. The oligomers were deprotected and cleaved from the solid support by treatment with aqueous ammonia (33%) overnight. After gel filtration on a NAP-25 column (Sephadex G25-DNA grade; Pharmacia) using water as eluent, the crude mixture was analyzed through a Mono-Q HR 5/5 anion exchange column, after which purification was achieved by using a Mono-Q HR 10/100 GL column (Pharmacia) with the following gradient systems: A = 20 mM Tris-HCl, 10 mM $NaClO_4$ in 15% CH_3CN, pH 7.4; B = 20 mM Tris-HCl, 600 mM $NaClO_4$ in 15% CH_3CN, pH 7.4. The low-pressure liquid chromatography system consisted of a Merck–Hitachi L-6200A intelligent pump, Mono-Q HR 10/100 GL column (Pharmacia), Uvicord SII 2138 UV detector (Pharmacia-LKB), and recorder. The product-containing fraction was desalted on a NAP-25 column and lyophilized. The presence of the desired product was confirmed by mass spectrometry.

UV Melting Experiments. Oligomers were dissolved in a buffer solution containing NaCl (0.1 or 1 M), potassium phosphate (0.02 M, pH 7.5), and EDTA (0.1 mM). The concentration was determined by measuring the absorbance in Milli-Q water at 260 nm and 80 °C, and assuming that the phosphonate nucleosides have the same extinction coefficients per base moiety in the denatured state as their natural nucleoside counterparts (A, ε = 15060; T, ε = 8560; C, ε = 7100; G, ε = 12180; T, ε = 8560; I, ε = 7500; U, ε = 9660). For all experiments, the concentration for each strand was 4 μM. Melting curves were determined with a Varian Cary 100 BIO spectrophotometer. Cuvettes were maintained at a 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 the Cary WinUV thermal application software. A quick heating and cooling cycle was carried out to allow proper annealing of both strands. The samples were then heated from 10 to 80 °C at a rate of 0.2 °C min^{-1} and were cooled again at the same speed. Melting temperatures were determined by plotting the first derivative of the absorbance as a function of temperature; data plotted were the average of two runs. Generally, up and down curves showed identical T_m values.

Primer Extensions. Primer extension reactions were typically carried out as 8 μl reactions with 0.2 μM primer tag01F3, 0.4 μM template TempN or TempNv3, 50 μM each PMTNpp, 0.2 μM TgoT_EPFLH in 1x Thermopol buffer. $MnCl_2$ was added to a final concentration of 0.6 mM for earlier polymerase variants. A typical thermal cycling program was 1 min at 94 °C, 5 min at 50 °C, and 2 h at 65 °C. For time courses, reaction mixtures were made up without polymerase and annealed by heating to 95 °C for 5 min, cooling to 4 °C at 0.1°C/s and held at 4 °C for 3-5 min. Polymerase was then added to the desired final concentration, typically from 20x stocks (i.e., 5 μl 10 μM polymerase stock added to 95 μl reaction mix for 0.5 μM final concentration). Reactions were aliquoted to 6 μl reactions in individual 0.2 ml tubes. These were placed simultaneously on a PCR block preheated to either 50 or 65 °C and tubes were removed from the heat and immediately quenched by addition of two volumes of 98% formamide and 10 mM EDTA at the desired timepoints. Primer extensions were analyzed by denaturing polyacrylamide gel electrophoresis (8 M urea, 16% acrylamide, 1x TBE). Large scale syntheses and purification are described in the Supplementary Information.

tPhoNA Reverse Transcription and PCR. Reverse transcription reactions for sequencing consisted of 1x Thermopol buffer (NEB), 0.25 μM primer outnest2_test7, 100 μM each dNTP, 30 nM T.RT521 or K.RT521K. A mastermix without template or polymerase was made up and divided into 17 μl aliquots. 2 μl gel-purified tPhoNA was added and primers annealed by heating to 5 min for 95 °C and cooling to 4 °C at 0.1 °C/s. After 3-5 min at 4 °C, 1 μl 600 nM polymerase was added (final concentration 30 nM) and incubated for 30 s at 50 °C, followed by 2 h at 65 °C. For RT gels, conditions were the same except primer IR700_test7 and template TempN2 were used and incubation was 5 min at 50 °C followed by 1 h at 65 °C.

After incubation, excess primers were removed by adding 0.5 μl (10U) exonuclease I (NEB) and incubating at 37 °C for 30 min. Reactions were then purified using Monarch PCR & DNA Cleanup kits (NEB), following the protocol for short PCR products.

PCR amplification was carried out using MyTaq HS DNA Polymerase (Bioline, London, UK) using primers outnest1 and outnest2 in reactions consisting 1x MyTaq buffer, 0.4 µM each primer, and 0.025 U/µl MyTaq DNA Polymerase. PCR reactions were typically run for 1 min at 98 °C, then 18 cycles of 15 s at 98 °C, 15 s at 55 °C, and 10 s at 72 °C, with a final 1 min at 68 °C polishing step.

Fourier Transform Infrared (FTIR) Spectroscopy. tPhoNA chimeras were synthesized in reactions consisting of 1x Thermopol buffer, 1 µM primer TagBspQI, 2 µM template TempN_MS2nick, 50 µM each PMTNpp, 0.1 mM $MnCl_2$, 1.3 µM TgotT_EPFLH and incubated in 24 x 40 µl aliquots in 0.2 ml 8-well PCR strips for 2 min at 94 °C, 10 min at 50 °C, and 4 h at 65 °C. DNA controls were synthesized in reactions consisting 1x MyTaq buffer, 1 µM primer TagBspQI, 1.8 µM template TempN_MS2nick, 0.01 U/µl My Taq HS DNA Polymerase (Bioline) and incubated for 2 min at 98 °C and 5 cycles of 15 s at 98 °C, 15 s at 50 °C, and 15 s at 72 °C. Purification was carried out as described in Supplementary Information for the large-scale tPhoNA synthesis and purification, except that the templates were 3'-biotinylated, meaning beads were washed in 2 x 100 µl 30 mM NaOH and the eluate contained the synthesized strand. An equivalent volume of formamide/EDTA loading buffer was added and the whole sample was loaded directly onto gels for purification. Yield was 287 pmol for tPhoNA and 192 pmol for DNA. Samples were washed 5 times in water using Vivaspin 500 centrifugal concentrators (Sartorius Stedim Biotech GmbH, Germany) with 10 kD MWCO.

Next, 10 µL of the resulting 10 µM solution of DNA or tPhoNA (prepared in double distilled water) were placed on a 3 mm diamond prism with KRS5 optics (1-reflection DuraSampl*IR* II, SensIR/Smith Detection). After drying with a stream of N_2 gas, IR spectra were recorded at 295 K and 4 cm^{-1} resolution in attenuated total reflectance (ATR) mode on a Bruker Vertex 80v spectrometer equipped with a liquid nitrogen-cooled MCT-C detector and a KBr beamsplitter. The optics compartment was kept under vacuum (<2 hPa) during data acquisition. All frequencies cited have an accuracy of ± 1 cm^{-1}. For each IR spectrum recorded, 500 interferograms were averaged before Fourier transformation. A reference spectrum of the clean ATR prism was recorded and used as a background for the sample spectra.

***In Vivo* Propagation Assay of DNA-PhoNA-DNA Chimeras.** Oligonucleotides were dissolved in Milli-Q water to reach 1 mM concentration and then diluted ten-fold before the assay. Oligonucleotides were tested inside a gapped heteroduplex generated through enzymatic modification, denaturation, and hybridization of the ampicillin resistance gene containing pAK1 and 2 plasmids.[43] The form of this heteroduplex is described in details elsewhere.[5,42] A mix of equimolar (25 ng each) purified NheI and NsiI cut pAK1, purified EcoRI cut, and dephosphorylated pAK 2 were diluted in 10 mM Tris-HCl pH 7.5 with 100 mM NaCl. The mixture was denaturated at 95 °C for 5 min before cooling to ambient temperature over 2 h to achieve hybridization, followed by water dialysis through a 0.05 µM nitrocellulose filter (Millipore) for 30 min. The oligonucleotides (100 pmoles) as well as a positive control (oligonucleotide with an intact catalytic residue codon) and two negative controls (water and oligomer with a deleted catalytic residue codon) were added separately to the dialyzed heteroduplex mixture in a 1x DNA ligase T4 reaction buffer (NEB) for 20 µL per sample. The mixture was then denaturated at 85 °C and cooled as before. Ligation was performed by adding 1 mM ATP and 400 U T4 DNA ligase (NEB) to the samples before overnight incubation at 16 °C. The ligated mixtures were dialyzed as before, and transformed by electroporation into a fresh electrocompetent *E. coli* K12 strain (Δ*thyA*:*aad*). Incubation of the electroporated bacteria was performed at 37 °C for 1 h, before plating 100 µL of a serial dilution of the suspension onto a Muller-Hinton (MH) media containing 100 µg mL^{-1} ampicillin (spreading the 10^0 and 10^{-1} dilutions) and onto the same media additionally supplemented with 0.3 mM thymidine (10^{-3} and 10^{-4} dilutions).

ASSOCIATED CONTENT

Supporting Information

The supporting information is available free of charge on the ACS Publications websites at DOI:

Detailed synthetic procedures, characterization information, NMR spectra, supplementary methods, figures, and references.

AUTHOR INFORMATION

Corresponding Author

*E-mail: piet.herdewijn@kuleuven.be; v.pinheiro@ucl.ac.uk

Notes

The authors declare no competing financial interest.

ACKNOWLEDGEMENTS

C.L. acknowledges the China Scholarship Council (CSC) for funding (grant 201306220065). The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Program (FP7/2007–2013)/ERC (Grant agreement no ERC-2012-ADG 20120216/ 320683). We also wish to thank FWO Vlaanderen (G0A8816N) and ERASynBio (*in vivo* XNA) for financial support. Mass spectrometry was conducted with the support of the Hercules Foundation of the Flemish Government (grant 20100225-7). We wish to express our gratitude to Guy Schepers, Luc Baudemprez, and Lia Margamuljana for excellent technical assistance. V.B.P. and C.C. were supported by the UK Biotechnology and Biosciences Research Council (grant number BB/N01023X/1). V.B.P. was also supported by the European Research Council [ERC-2013-StG project 336936 (HNAepisome)]. A.M. was supported by the Medical Research Council UK (Career Development Award MR/M00936X/1).

REFERENCES

(1) Herdewijn, P.; Marliere, P. *Chem. Biodiv.* **2009**, *6*, 791-808.
(2) Chaput, J. C.; Yu, H. Y.; Zhang, S. *Chem. Biol.* **2012**, *19*, 1360-1371.
(3) Larsen, A. C.; Dunn, M. R.; Hatch, A.; Sau, S. P.; Youngbull, C.; Chaput, J. C. *Nat. Commun.* **2016**, *7*, 11235.
(4) Pinheiro, V. B.; Taylor, A. I.; Cozens, C.; Abramov, M.; Renders, M.; Zhang, S.; Chaput, J. C.; Wengel, J.; Peak-Chew, S. Y.; McLaughlin, S. H.; Herdewijn, P.; Holliger, P. *Science* **2012**, *336*, 341-344.
(5) Pezo, V.; Liu, F. W.; Abramov, M.; Froeyen, M.; Herdewijn, P.; Marliere, P. *Angew. Chem. Int. Ed.* **2013**, *52*, 8139-8143.
(6) Bande, O.; Abu El Asrar, R.; Braddick, D.; Dumbre, S.; Pezo, V.; Schepers, G.; Pinheiro, V. B.; Lescrinier, E.; Holliger, P.; Marliere, P.; Herdewijn, P. *Chem. Eur. J.* **2015**, *21*, 5009-5022.

(7) Bande, O.; Braddick, D.; Agnello, S.; Jang, M.; Pezo, V.; Schepers, G.; Rozenski, J.; Lescrinier, E.; Marlière, P.; Herdewijn, P. *Chem. Sci.* **2016**, *7*, 995-1010.

(8) Pezo, V.; Schepers, G.; Lambertucci, C.; Marlière, P.; Herdewijn, P. *ChemBioChem* **2014**, *15*, 2255-2258.

(9) Inoue, N.; Shionoya, A.; Minakawa, N.; Kawakami, A.; Ogawa, N.; Matsuda, A. *J. Am. Chem. Soc.* **2007**, *129*, 15424-15425.

(10) Marlière, P.; Patrouix, J.; Döring, V.; Herdewijn, P.; Tricot, S.; Cruveiller, S.; Bouzon, M.; Mutzel, R. *Angew. Chem. Int. Ed.* **2011**, *50*, 7109-7114.

(11) Mehta, A. P.; Li, H.; Reed, S. A.; Supekova, L.; Javahishvili, T.; Schultz, P. G. *J. Am. Chem. Soc.* **2016**, *138*, 14230-14233.

(12) Eremeeva, E.; Abramov, M.; Margamuljana, L.; Rozenski, J.; Pezo, V.; Marliere, P.; Herdewijn, P. *Angew. Chem. Int. Ed.* **2016**, *55*, 7515-7519.

(13) Zhang, Y.; Ptacin, J. L.; Fischer, E. C.; Aerni, H. R.; Caffaro, C. E.; San Jose, K.; Feldman, A. W.; Turner, C. R.; Romesberg, F. E. *Nature* **2017**, *551*, 644-647.

(14) Schoning, K. U.; Scholz, P.; Guntha, S.; Wu, X.; Krishnamurthy, R.; Eschenmoser, A. *Science* **2000**, *290*, 1347-1351.

(15) Yu, H. Y.; Zhang, S.; Chaput, J. C. *Nat. Chem.* **2012**, *4*, 183-187.

(16) Schoning, K. U.; Scholz, P.; Wu, X. L.; Guntha, S.; Delgado, G.; Krishnamurthy, R.; Eschenmoser, A. *Helv. Chim. Acta* **2002**, *85*, 4111-4153.

(17) Ichida, J. K.; Horhota, A.; Zou, K. Y.; McLaughlin, L. W.; Szostak, J. W. *Nucleic Acids Res.* **2005**, *33*, 5219-5225.

(18) Yu, H. Y.; Zhang, S.; Dunn, M. R.; Chaput, J. C. *J. Am. Chem. Soc.* **2013**, *135*, 3583-3591.

(19) Dunn, M. R.; Chaput, J. C. *ChemBioChem* **2016**, *17*, 1804-1808.

(20) Renders, M.; Emmerechts, G.; Rozenski, J.; Krecmerova, M.; Holy, A.; Herdewijn, P. *Angew. Chem. Int. Ed.* **2007**, *46*, 2501-2504.

(21) Renders, M.; Lievrouw, R.; Krecmerova, M.; Holy, A.; Herdewijn, P. *ChemBioChem* **2008**, *9*, 2883-2888.

(22) Rejman, D.; Snasel, J.; Liboska, R.; Tocik, Z.; Paces, O.; Kralikova, S.; Rinnova, M.; Kois, P.; Rosenberg, I. *Nucleosides Nucleotides Nucleic Acids* **2001**, *20*, 819-823.

(23) Sipova, H.; Springer, T.; Rejman, D.; Simak, O.; Petrova, M.; Novak, P.; Rosenbergova, S.; Pav, O.; Liboska, R.; Barvik, I.; Stepanek, J.; Rosenberg, I.; Homola, J. *Nucleic Acids Res.* **2014**, *42*, 5378-5389.

(24) Liu, C.; Dumbre, S. G.; Pannecouque, C.; Huang, C. S.; Ptak, R. G.; Murray, M. G.; De Jonghe, S.; Herdewijn, P. *J. Med. Chem.* **2016**, *59*, 9513-9531.

(25) Liu, C.; Dumbre, S. G.; Pannecouque, C.; Korba, B.; De Jonghe, S.; Herdewijn, P. *Org. Biomol. Chem.* **2017**, *15*, 5513-5528.

(26) Kostov, O.; Pav, O.; Budesinsky, M.; Liboska, R.; Simak, O.; Petrova, M.; Novak, P.; Rosenberg, I. *Org. Lett.* **2016**, *18*, 2704-2707.

(27) Wu, T. F.; Froeyen, M.; Kempeneers, V.; Pannecouque, C.; Wang, J.; Busson, R.; De Clercq, E.; Herdewijn, P. *J. Am. Chem. Soc.* **2005**, *127*, 5056-5065.

(28) Moffatt, J. G.; Khorana, H. G. *J. Am. Chem. Soc.* **1961**, *83*, 649-658.

(29) Lesnik, E. A.; Freier, S. M. *Biochemistry* **1995**, *34*, 10807-10815.

(30) Pav, O.; Kosiova, I.; Barvik, I.; Pohl, R.; Budesinsky, M.; Rosenberg, I. *Org. Biomol. Chem.* **2011**, *9*, 6120-6126.

(31) Pinheiro, V. B.; Loakes, D.; Holliger, P. *Bioessays* **2013**, *35*, 113-122.

(32) Fogg, M. J.; Pearl, L. H.; Connolly, B. A. *Nat. Struct. Biol.* **2002**, *9*, 922-927.

(33) Gardner, A. F.; Jack, W. E. *Nucleic Acids Res.* **2002**, *30*, 605-613.

(34) Houlihan, G.; Arangundy-Franklin, S.; Holliger, P. *Curr. Opin. Biotechnol.* **2017**, *48*, 168-179.

(35) Kuroita, T.; Matsumura, H.; Yokota, N.; Kitabayashi, M.; Hashimoto, H.; Inoue, T.; Imanaka, T.; Kai, Y. *J. Mol. Biol.* **2005**, *351*, 291-298.

(36) Cozens, C.; Pinheiro, V. B.; Vaisman, A.; Woodgate, R.; Holliger, P. *Proc. Nat. Acad. Sci. U S A* **2012**, *109*, 8067-8072.

(37) Cozens, C.; Mutschler, H.; Nelson, G. M.; Houlihan, G.; Taylor, A. I.; Holliger, P. *Angew. Chem. Int. Ed.* **2015**, *54*, 15570-15573.

(38) Taylor, A. I.; Pinheiro, V. B.; Smola, M. J.; Morgunov, A. S.; Peak-Chew, S.; Cozens, C.; Weeks, K. M.; Herdewijn, P.; Holliger, P. *Nature* **2015**, *518*, 427-430.

(39) Kropp, H. M.; Betz, K.; Wirth, J.; Diederichs, K.; Marx, A. *PLoS One* **2017**, *12*, 1-20.

(40) Chim, N.; Shi, C. H.; Sau, S. P.; Nikoomanzar, A.; Chaput, J. C. *Nat. Commun.* **2017**, *8*, 1810.

(41) Tsai, C. H.; Chen, J. Y.; Szostak, J. W. *Proc. Nat. Acad. Sci. U S A* **2007**, *104*, 14598-14603.

(42) Pochet, S.; Kaminski, P. A.; Van Aerschot, A.; Herdewijn, P.; Marliere, P. *C. R. Biol.* **2003**, *326*, 1175-1184.

(43) Tchou, J.; Kasai, H.; Shibutani, S.; Chung, M. H.; Laval, J.; Grollman, A. P.; Nishimura, S. *Proc. Natl. Acad. Sci. U. S. A.* **1991**, *88*, 4690-4694.

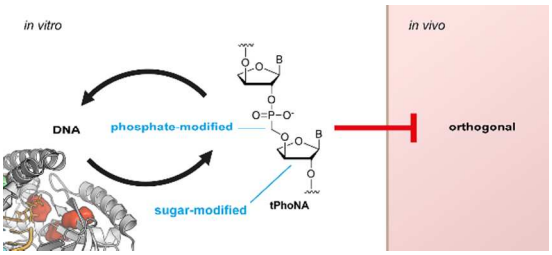
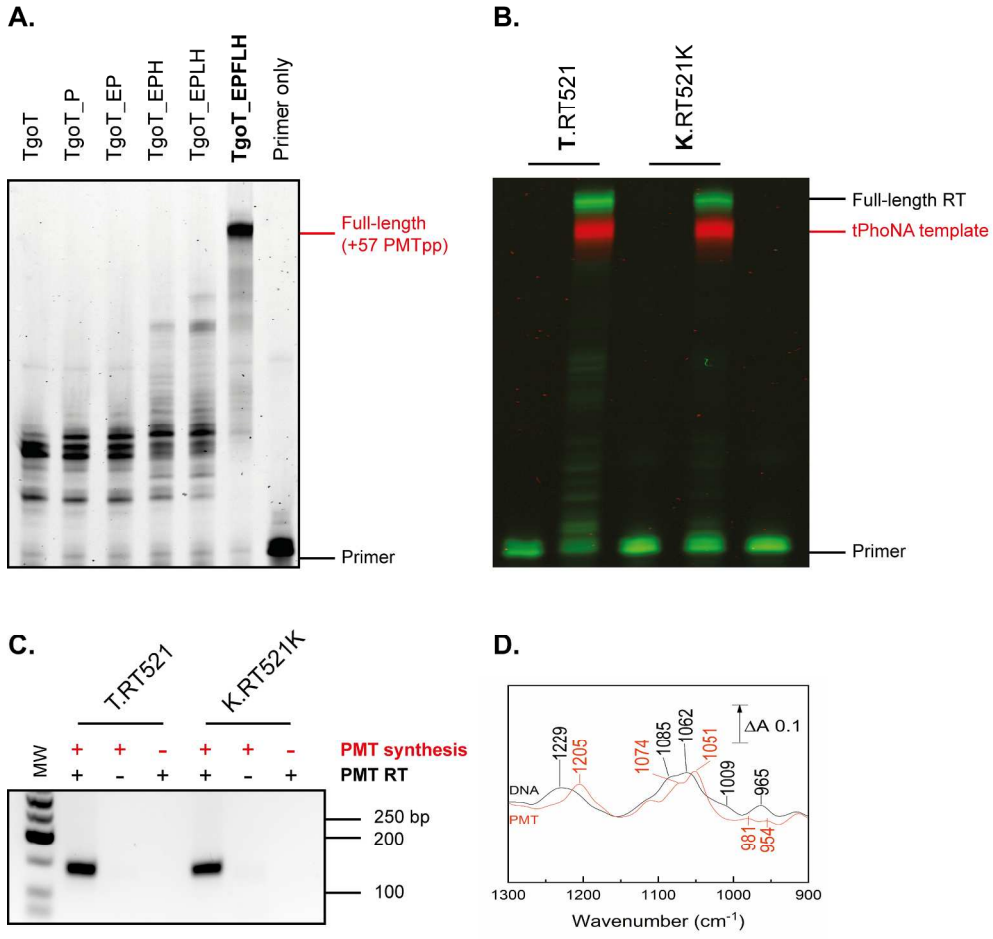
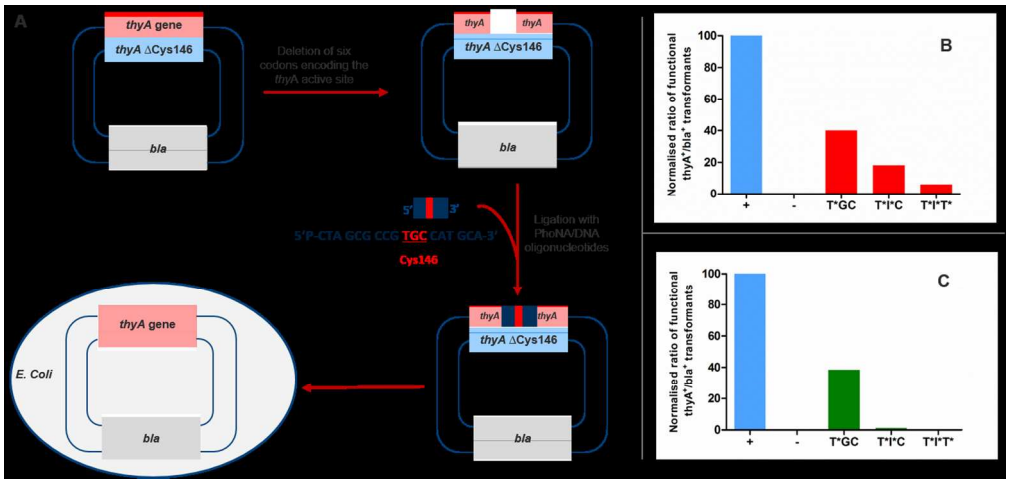


Table of Contents



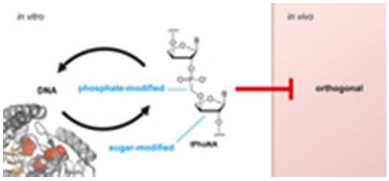
tPhoNA as a genetic polymer. (A) Impact of the systematic engineering of the starting polymerase (TgoT) for the synthesis of tPhoNA under forcing conditions (Figure S5 for details). Intermediates in the engineering process including TgoT_P (TgoT: L403P), TgoT_EP (TgoT_P: H147E), Tgo_EPH (TgoT_EP: E664H), TgoT_EPLH (TgoT_EPH: I521L) as well as the isolated tPhoNA synthase TgoT_EPFLH (TgoEPH: L408F) are shown. (B) Reverse transcription of tPhoNA to DNA. Both T.RT521 and K.RT521K were capable of template-dependent DNA synthesis (green) against a chimeric tPhoNA template (20 nt DNA from the original synthesis primer with 39 PMT incor-porations based on TempN2), which is shown in red, from an RNA primer. (C) PCR amplification of the tPhoNA RT products was only possible when both tPhoNA synthesis and reverse transcription were carried out, confirming that tPhoNA can store and propagate genetic information. (MW: NEB Low molecular weight DNA ladder). (D) ATR-FTIR spectra of DNA (black) and tPhoNA (red) in the 1300-900 cm^{-1} range. Frequency changes of the (PO_2^-) ν_{as} from 1225-1240 to 1205 cm^{-1}, a shift of the ($PO_2^{sup>-}$) symmetric stretching mode (ν_s) at 1085 cm^{-1}, and the (C-O) ribose stretch at 1062 cm^{-1} in DNA to 1074 and 1051 cm^{-1} in tPhoNA can be observed as well as signal reductions at 1009 cm^{-1} (assigned to a ribose ring mode) and the characteristic ribose-phosphodiester skeletal motion mode of DNA at 965 cm^{-1}. Com-plete spectra are shown in Figure S10.

375x355mm (168 x 168 DPI)



(A) Selection screening of PhoNA containing oligonucleotides capable of templating DNA synthesis in *E. coli* based on the restoration of the *thyA* gene.5,42 *In vivo* propagation of DNA-tPhoNA-DNA (B) and DNA-dPhoNA-DNA (C) chimeras. T* and I* indicate Pho-modified units with a thymine and hypoxan-thine base, respectively. The normalized ratios correspond to the experimentally derived average number of thymidine prototrophic colonies (*thyA*+) from the average total number of colonies (*bla*+).

378x179mm (96 x 96 DPI)



TOC

16x7mm (300 x 300 DPI)