Invading *Escherichia coli* Genetics with a Xenobiotic Nucleic Acid Carrying an Acyclic Phosphonate Backbone (ZNA)

Min Luo,[§] Elisabetta Groaz,[§] Mathy Froeyen,[§] Valérie Pezo,[‡] Faten Jaziri,[‡,†] Piotr Leonczak,[§] Guy Schepers,[§] Jef Rozenski,[§] Philippe Marlière,[‡] and Piet Herdewijn[§,*]

[§] KU Leuven, Rega Institute for Medical Research, Medicinal Chemistry, Herestraat 49 – box 1041, 3000 Leuven, Belgium

[‡] Génomique Métabolique, Genoscope, Institut François Jacob, CEA, CNRS, Univ Evry, Université Paris-Saclay, 2 Rue Gaston Crémieux 91057 Evry, France

ABSTRACT: A synthetic orthogonal polymer embracing a chiral acyclic-phosphonate backbone [(*S*)-ZNA] is presented that uniquely adds to the emerging family of xenobiotic nucleic acids (XNAs). (*S*)-ZNA consists of reiterating six-atom structural units and can be accessed in few synthetic steps from readily available phophonomethylglycerol nucleoside (PMGN) precursors. Comparative thermal stability experiments conducted on homo- and heteroduplexes made of (*S*)-ZNA are described that evince its high self-hybridization efficiency in contrast to poor binding of natural complements. Although preliminary and not conclusive, circular dichroism data and dynamic modeling computations provide support to a left-handed geometry of double-stranded (*S*)-ZNA. Nonetheless, PMGN diphosphate monomers were recognized as substrates by *Escherichia coli* (*E. coli)* polymerase I as well as being imported into *E. coli* cells equipped with an algal nucleotide transporter. A further investigation into the *in vivo* propagation of (*S*)-ZNA culminated with the demonstration of the first synthetic nucleic acid with an acyclic backbone that can be transliterated to DNA by the *E. coli* cellular machinery.

INTRODUCTION

The development of synthetic genetic systems capable of undergoing evolution to deliver functional polymers as well as encoding novel traits in living cells is driven by the necessity to expand current biotechnologies as well as either generating or optimizing existing cellular processes carried out by nucleic acids for various purposes.[1-2] Overall, this is a very complex task that involves the careful planning and chemical synthesis of naturally unavailable or xenobiotic nucleic acids (XNAs), engineering of enzymes that can polymerize their precursors with high efficiency and fidelity, and ultimately the *in vivo* enforcement of both XNAs and evolved polymerases.[3] This last step can be particularly challenging as multiple criteria need to be met in the selection of an XNA polymer for *in vivo* applications. Ideally, the modified backbone should be metabolically accessible by the action of cellular polymerases from suitably activated monomers. To this end, we need the ability to chemically manufacture XNA precursors with minimal synthetic effort, which can obviate the need of inefficient intracellular conversion to enzymatically unstable monophosphorylated species. Furthermore, an XNA system should form relatively stable homoduplexes in order to induce folding processes, while fulfilling the orthogonality condition of hybridizing with poor efficiency with natural DNA and RNA.

To date, a few synthetic elements have been implanted *in vivo*, and these are mostly limited to base-modified DNA. The reconstitution of replication, transcription, and translation *in vivo* has been demonstrated at the level of a single base-pair inserted in a plasmid.[4-5] The *Escherichia coli* (*E. coli)* cellular machinery was also proven to sustain self-reproduction upon replacement of single synthetic nucleobases throughout its entire genome, while leaving the three other nucleobases untouched.[6-8] However, all four nucleobases could be exchanged with artificial bases within a single gene and transliterated into DNA *in vivo*.[9] Short stretches of DNA containing size-expanded analogues of natural nucleobases were shown to be converted into DNA and propagate in living *E. coli* cells.[10] DNA with a few triazole linkers in its backbone in the place of a phosphodiester bond has been accepted by *E.coli*[11-12] and human cells.[13] In the case of sugar modified XNAs, the transfer of genetic information from DNA to a specific XNA architecture and back into DNA was demonstrated to occur *in vitro*.[14] *In vivo*, a four codon XNA inserted in a plasmid could be transliterated into DNA and give rise to a functional protein in *E. coli*,[15] while 4'-thio-DNA could be transcribed into RNA in mammalian cells.[16-17]

Currently, the general utility of many of these chemistries is largely limited by difficulties to achieve orthogonality, high probability of causing mutations, poor metabolic stability as well as inability to generate metabolic precursors *in vivo*. For these reasons and in order to advance the tool-box for xenobiology, we have devised and investigated an XNA analogue with an acyclic methylphosphonate backbone, which was named ZNA (Figure 1). This molecule can exist in two enantiomeric forms, i.e., (*S*)-and (*R*)-ZNA, and is made up of repeating phosphonomethylglycerol (PMG) units, which are structurally identical to the hydroxy(phosphonomethoxy)propyl motif introduced by Holý in the antiviral field.[18] The phosphonomethylglycerol nucleoside (PMGN) monomers are connected through a system of six-bonds in analogy to DNA and unlike the closest phosphate congener GNA,[19] which is based on a five-bond internucleotide linkage. Recently, we demonstrated that the phosphonate motif may represent a convenient modification for

evolving an orthogonal polymerase that poorly recognises natural substrates.[20]

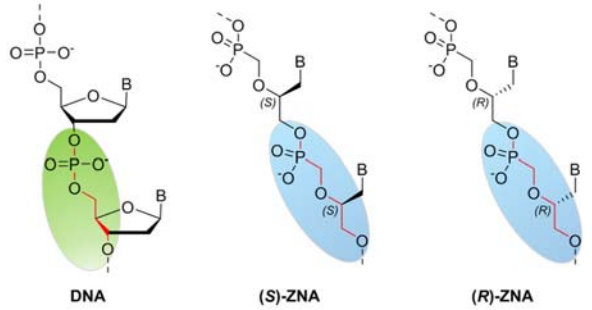


Figure 1. Comparison between the phosphodiester and phosphonomethylester backbone of DNA and enantiomeric (*S*)- and (*R*)-ZNA. All structures share a repeating six-atom internucleotide linkage.

RESULTS AND DISCUSSION

Hybridization Properties of (*S*)- and (*R*)-ZNA. *Chemical Synthesis of Phosphonate-Linked Acyclic Chimeric Oligonucleotides*. The synthesis of oligomers containing a chiral acyclic sugar-like phosphonate backbone was achieved through a combination of different solid-phase chemistries. In the case of chimeric sequences alternating phosphodiester and phosphonate internucleotide linkages, the standard phosphoramidite approach was used for coupling between natural protected monomers, while the introduction of a single PMGN residue was preferably achieved using a modified version of the *H*-phosphonate method (Method A, Table S1). Fully modified oligonucleotides were synthesized according to a further adapted *H*-phosphonate protocol (Method B, Table S1) starting from the *H*-phosphinate monomers **1a**,**b**–**4a**,**b** shown in Figure 2.

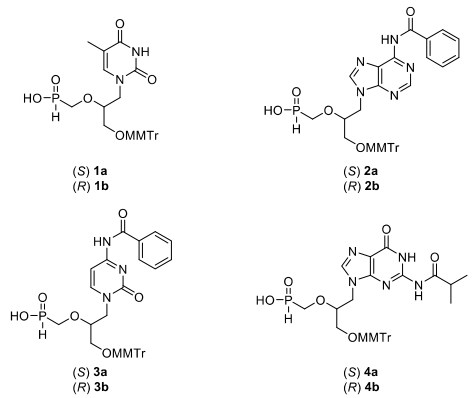


Figure 2. Chemical structures of the (*S*)- and (*R*)-phosphonomethylglycerol nucleoside (PMGN) *H*-phosphinate building blocks used for the synthesis of enantiomeric phosphonate-linked acyclic oligonucleotides.

A customized phosphoramidite protocol using PMGN phosphonamidites **S29** and **S30** (Scheme S6)[21] as building blocks was also employed for the synthesis of two oligonucleotide chimeras (**ON11**, **ON12**, Table S1 and S2), however this method was later discontinued due to the difficulties encountered in the synthesis of sufficient amounts of amidite monomers. It should be mentioned that the oligonucleotide synthesis was performed in the 5' to 3' direction instead of the classical 3' to 5' direction. The detailed reactions conditions for each method as well as the synthetic procedures used for the preparation of monomers **1a**,**b**–**4a**,**b** are reported in the Supporting Information (Table S1, Figure S1, Schemes S1-S4). The oligonucleotides sequences used in this study and their mass analysis are listed in Table S2 (Supporting Information).

Thermal Stability of ZNA Hetero- and Homoduplexes. In order to investigate the hybridization properties of a series of oligonucleotide chimeras containing PMGN building block modifications, the melting temperatures (T_m) of the resulting duplexes were determined by temperature-dependent UV spectroscopy (Table 1). It was observed that the T_m of the chimera/DNA duplexes was lowered by at least -7.5 °C upon introduction of either a single (*S*)- or (*R*)-PMGN building block at the center of the strand or in both strands (Table 1, entries 5-12). No significant difference in the destabilizing effect was observed between the (*S*)- and (*R*)-enantiomer. These T_m measurements revealed an instability effect due to the introduction of the acyclic PMGN monomers, in agreement with previous reports describing a similar effect for oligonucleotide duplexes containing flexible nucleotides.[22] Notably, the T_m of the fully modified (*S*)-ZNA duplex was higher (37.3 °C, Table 1, entry 15) than that measured for the corresponding natural DNA duplex (32.8 °C, Table 1, entry 1). On the other hand, no hybridization was observed between (*S*)-ZNA and DNA or RNA for A/T containing sequences (Table 1, entries 13, 14). As this property, i.e., excellent self-hybridization and poor hybridization with DNA and RNA, is important for the aim of our project, we synthesized two sequences for each enantiomeric series bearing the four canonical bases A, G, C, and T. In all cases (**ON7**–**ON10**, Table S1), a DNA nucleotide was present at the 3'- and 5'-end since acyclic oligonucleotides with a terminal free hydroxyl group tend to degrade by undergoing a transesterification reaction under strong basic deprotection conditions. Furthermore, having a natural nucleotide at the 5'-end simplifies oligonucleotide synthesis. T_m measurements of the fully modified (*S*)-ZNA duplex confirmed the initial results as a stronger self-hybridization was observed for (*S*)-dsZNA (57.7 °C, Table 1, entry 16) compared to dsDNA (48.8 °C, Table 1, entry 2), and even stronger than for (*S*)-ZNA:DNA duplexes (39.4 °C entry 17, and 41.4 °C, entry 18, Table 1). No or very week hybridization with RNA was observed. The melting curves for (*S*)-ZNA are shown in Figure 3. Remarkably, T_m studies conducted on the corresponding enantiomeric sequences revealed that (*R*)-ZNA did not hybridize with DNA or RNA (Table 1, entries 20 and 21) while as expected also showing a stronger self-hybridization ability than DNA (57.2 versus 48.8 °C, Table 1, entries 19 and 1, respectively). This property may be advantageous for XNA folding to avoid strand displacement with DNA or RNA.

Table 1. Tm values of non-self-complementary antiparallel-stranded oligonucleotide duplexes.[(a)]

Entry	Duplex	Sequence	T_m [°C]
1	DNA1:DNA2	5′-TTATATTTAAAA-3′/3′-AATATAAATTTT-5′	32.8
2	DNA3:DNA4	5′-TGCATCAGTA-3′/3′-ACGTAGTCAT-5′	48.8
3	RNA1:RNA2	5′-UUAUAUUUAAAA-3′/3′-AAUAUAAAUUUU-5′	31.7
4	RNA3:RNA4	5′-UGCAUCAGUA-3′/3′-ACGUAGUCAU-5′	59.9
5	**ON1**:DNA2	(*S*)-5′-TTATA**t**TTAAAA-3′/3′-AATATAAATTTT-5′	18.9
6	**ON2**:DNA2	(*R*)-5′-TTATA**t**TTAAAA-3′/3′-AATATAAATTTT-5′	17.3
7	**ON3**:DNA1	(*S*)-5′-TTTTAA**a**TATAA-3′/3′-AAAATTTATATT-5′	25.3
8	**ON4**:DNA1	(*R*)-5′-TTTTAA**a**TATAA-3′/3′-AAAATTTATATT-5′	23.5
9	**ON1**:**ON3**	(*S*)-5′-TTATA**t**TTAAAA-3′/(*S*)-3′-AATAT**a**AATTTT-5′	< 10
10	**ON2**:**ON4**	(*R*)-5′-TTATA**t**TTAAAA-3′/(*R*)-3′-AATAT**a**AATTTT-5′	< 10
11	**ON1**:**ON4**	(*S*)-5′-TTATA**t**TTAAAA-3′/(*R*)-3′-AATAT**a**AATTTT-5′	< 10
12	**ON2**:**ON3**	(*R*)-5′-TTATA**t**TTAAAA-3′/(*S*)-3′-AATAT**a**AATTTT-5′	< 10
13	**ON5**:DNA1	(*S*)-5′-T**tttaaatata**A-3′/3′-AAAATTTATATT-5′[(b)]	< 7
14	**ON6**:DNA2	(*S*)-5′-T**tatatttaaa**A-3′/3′-AATATAAATTTT-5′[(b)]	< 7
15	**ON5**:**ON6**	(*S*)-5′-T**tttaaatata**A-3′/(*S*)-3′-A**aaatttatat**T-5′	37.3
16	**ON7**:**ON8**	(*S*)-5′-T**gcatcagt**A-3′/(*S*)-3′-A**cgtagtca**T-5′	57.7
17	**ON7**:DNA4	(*S*)-5’-T**gcatcagt**A-3′/3′- ACGTAGTCAT-5′[(c)]	39.4
18	**ON8**:DNA3	(*S*)-5′-T**actgatgc**A-3′/3′-ATGACTACGT-5′[(d)]	41.4
19	**ON9**:**ON10**	(*R*)-5’-T**gcatcagt**A-3′/(*R*)-3’-A**cgtagtca**T-5′	57.2
20	**ON9**:DNA4	(*R*)-5’-T**gcatcagt**A-3′/3′-ACGTAGTCAT-5′[(e)]	–
21	**ON10**:DNA3	(*R*)-5′-T**actgatgc**A-3′/3′-ATGACTACGT-5′[(e)]	–

[(a)] Modified units are indicated in bold font, small letters. (*S*)- and (*R*)-ZNA nucleotides are shown as lowercase and lowercase underlines letters, respectively.[(b)] No Tm (< 3 °C) with RNA complement (U base instead of T base) was observed for **ON5**:RNA1 and **ON6**:RNA2; [(c)] Tm with the RNA complement for **ON7**:RNA4 is 14.1 °C (U base instead of T base). [(d)] No Tm (< 5 °C) with the RNA complement for **ON8**:RNA3 (U base instead of T base). [(e)] No Tm with the RNA complement (U base instead of T base) was observed for **ON9**:RNA4 and **ON10**:RNA3.

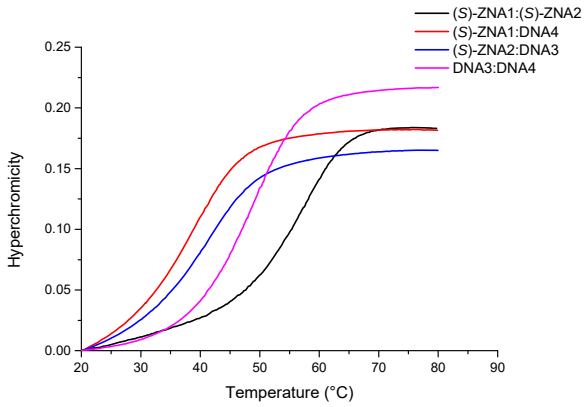


Figure 3. Melting curves of fully modified (*S*)-ZNA and hybrid (*S*)-ZNA:DNA duplexes as compared with DNA; [(*S*)-ZNA1 = **ON7**, (*S*)-ZNA2 = **ON8**].

Structure Determination of (*S*)- and (*R*)-ZNA Duplexes. *Circular Dichroism (CD) Measurements.* Due to the difference in hybridization properties between (*S*)- and (*R*)-ZNA, we investigated the structure of different hybrids by circular dichroism (CD) spectroscopy. The CD spectra of (*S*)-dsZNA, (*R*)-dsZNA, ON1/DNA2, ON2/DNA2, ON3/DNA1, and ON4/DNA1 duplexes were shown in Figures 4 and S2. As shown in Figure 4, the CD spectrum of double stranded (*R*)-ZNA was characterized by a maximum at 272 nm and a minimum approximately located at 242 nm. The (*R*)-ZNA signal appeared essentially as a mirror image of that arising from (*S*)-ZNA with a peak of great intensity at 268 nm and a small peak at 242 nm. Similar CD spectra were obtained with (*S*)-dsGNA and (*R*)-dsGNA.[19] The spectrum of natural dsDNA resembled that of (*R*)-ZNA in terms of shape although varying in peak position and intensity. The positive and negative CD bands at 278 and 253 nm, respectively, are diagnostic for a B-type helix.

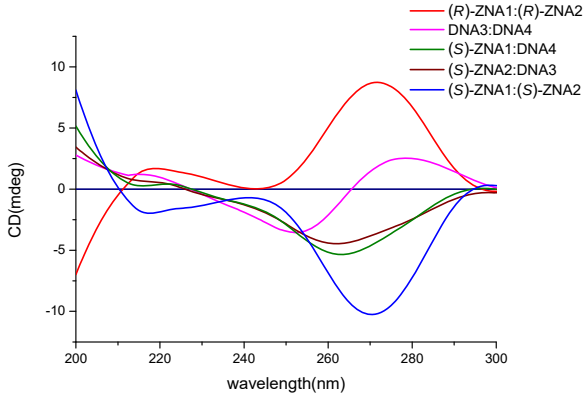


Figure 4. CD spectra of fully modified and hybrid ZNA duplexes in comparison to natural dsDNA; spectra were measured at 25 °C and are shown in units of molar ellipticity as a function of wavelength; [(*R*)-ZNA1 = **ON9**, (*R*)-ZNA2 = **ON10**, (*S*)-ZNA1 = **ON7**, (*S*)-ZNA2 = **ON8**].

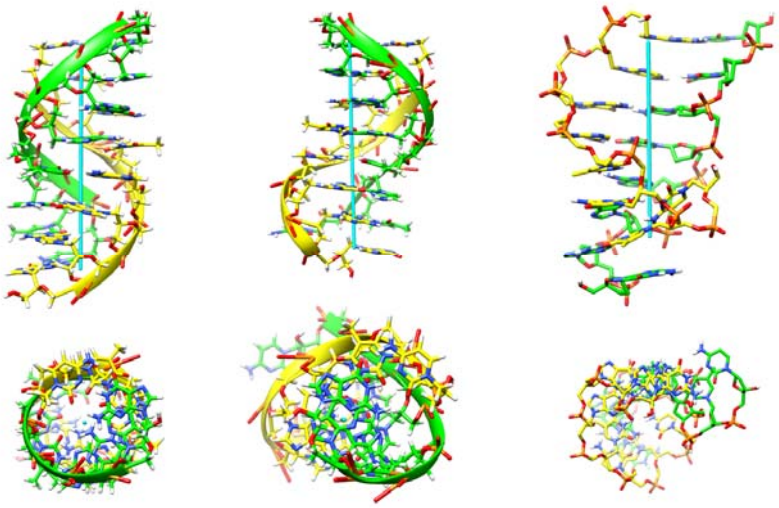
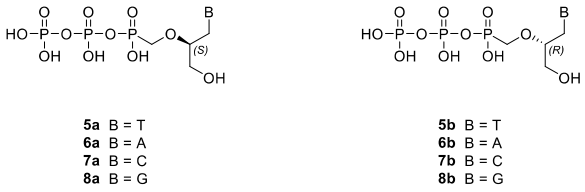


Figure 5. Representative conformations of (*R*)-ZNA:(*R*)-ZNA, (*S*)-ZNA:(*S*)-ZNA, and (*S*)-ZNA:DNA duplexes (side and top view). The (*S*)-ZNA:DNA simulated structure suffers from base pair fraying at the ends.[23-24]

Therefore, by comparing the different CD spectral shapes in Figure 4, it could be suggested that (*R*)-dsZNA favored the formation of a right-handed helix, while (*S*)-dsZNA adopted an uncommon left-handed double helical structure. However, X-ray crystallography analysis of (*S*)-dsGNA shows it to be a right-handed helix.[25-26] So, a definitive conclusion about the handedness of (*S*)-dsZNA and (*R*)-dsZNA in solution and in solid phase cannot be taken at this moments. More physicochemical studies are needed. The spectra of (*S*)-ZNA/DNA duplexes consisted of more flattened curves making it difficult to deduce the shape of these helixes.

Dynamic Modeling Study. To better understand the observed hybridization properties and further investigate the helical conformation assumed by the enantiomeric ZNA duplex structures, we built a dynamic model of these helixes. In particular, 8-mer (*S*)-dsZNA, (*R*)-dsZNA, (S)-ZNA/DNA, and (*R*)-ZNA/DNA duplexes were build starting from a classical 8-mer DNA duplex in a water box adding 14 Na^+ counter ions and molecular dynamics were continued for 100 ns. The (*R*)-ZNA/DNA duplex was found to be unstable, which agrees with the Tm measurements. Both (*S*)-dsZNA as (*R*)-dsZNA gave similar duplexes with opposite twist values, i.e., left- and right-handed, respectively. The abolition of the stereoelectronic effect due to the replacement of the oxygen atom of one phosphoester function by a methylene group increased the conformational flexibility of the phosphonate backbone.[27] Most of the O-P-CH_2(O) bonds, however, remained in the gauche conformation. The (*S*)-ZNA/DNA duplex is a remarkable structure as an example of a potential left-handed DNA coil with one base as repeating unit (Figure 5). The analysis of the backbone dihedral angles indicated large deviations (especially for α and ζ, less for β, δ, and ε) from the classical backbone angles as found in B-DNA. The lower Tm (Table 1) of the (*S*)-ZNA/DNA duplex could be attributed to this anomalous backbone combination not found in regular dsDNA.

***In vitro* and *in vivo* Performance of ZNA.** To determine the eligibility of ZNA as artificial genetic system in view of further *in vitro* and *in vivo* evolution, we conducted different studies. Firstly, we investigated the potential of PMGN modified triphosphate mimics to act as substrates for DNA elongation by *E. coli* replicative enzymes *in vitro*. For this purpose, it was necessary to synthesize two enantiomeric series of diphosphates of acyclic phosphonomethylglycerol nucleosides (PMGNpp), i.e., compounds **5a**–**8a** and **5b**–**8b** (Figure 6) with a (*S*) and (*R*) configuration, respectively (see Supplementary Information for full details concerning the synthesis and characterization of these compounds including Scheme S5). Among the five distinct DNA polymerases encoded by *E. coli*,[28] we selected for this study DNA Pol III and the Klenow fragment (3',5' exo^-) of *E. coli* DNA polymerase I, which have been recognized as the primary and most abundant replicases in *E. coli* cells, respectively. In addition, Vent (exo-) polymerase was selected as an example of a thermophilic DNA polymerase.



5a B = T
6a B = A
7a B = C
8a B = G

5b B = T
6b B = A
7b B = C
8b B = G

Figure 6. Chemical structures of diphosphates of (*S*)- and (*R*)-phosphonomethylglyceryl nucleosides (PMGNpp) investigated as substrates for polymerases *in vitro* and *in vivo*.

The efficiency of these polymerases to extend a DNA primer annealed to a DNA template upon incorporation of PMGNpp building blocks was evaluated by polyacrylamide gel-based single-nucleotide primer extension assays. An overview of the primer-template duplexes used in this study is given in Table S10 (Supporting Information). Although full-length products could not be obtained, it was most encouraging to observe that the Klenow (exo^-) DNA polymerase and Vent (exo-) polymerase could recognize and accept as substrates both (*S*)- and (*R*)-PMGNpp enantiomers containing the four canonical nucleobases (Figures 7 and S4). These results are fairly similar to those described for (*S*)- and (*R*)-GNA.[29-30] The enantiospecificity of the incorporation did not differ significantly, although slightly better substrate properties of the (*R*) counterparts for the Klenow fragment were observed, especially in the case of PMGApp and PMGCpp derivatives. Slight higher incorporation efficiencies were observed for the

(*S*) enantiomers in case of Vent (exo-) polymerase. In contrast, DNA IIIα polymerase barely succeeded to extend the primer by one residue in the presence of **5a**-**8a** and **5b**-**8b** (Figure S5). Thus, it could be assumed that Eco polymerase I could be a better candidate than Eco polymerase III for *in vitro* and *in vivo* evolution.

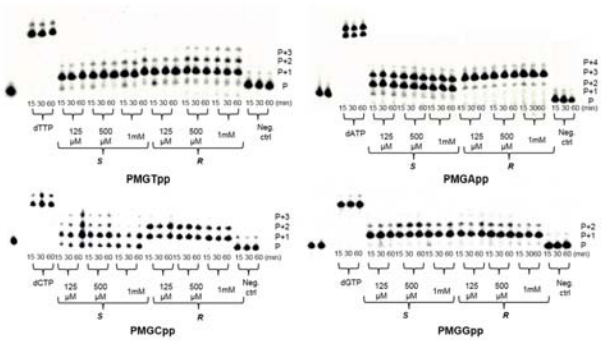


Figure 7. Profiles of incorporation of the dNTP controls (dTTP, dATP, dCTP, and dGTP), (*S*)/(*R*)-PMGTpp, (*S*)/(*R*)-PMGApp, (*S*)/(*R*)-PMGCpp, and (*S*)/(*R*)-PMGGpp by the Klenow (*exo*$^-$) fragment DNA polymerase (0.05 U.μL^{-1}) with Primer 1 and Template 1-4 (Table S10).

The specificity of PMGNpp building blocks incorporation by Klenow fragment DNA polymerase was assessed by primer extension experiments with template 5 (Table S10). Matched and mismatched pairs between purine and pyrimidine bases (A:(*S*)-PMGTpp, A:(*R*)-PMGTpp, A:(*S*)-PMGCpp, and A: (*R*)-PMGCpp) were compared. The results showed that for both enantiomeric PMGNpps, the thymine nucleoside phosphonate was incorporated versus adenine in the template, while almost no mismatched (*S*)- and (*R*)-PMGCpp enantiomers were incorporated (Figure S6).

Having established DNA dependent ZNA polymerase activity for Klenow (exo$^-$) DNA polymerase (Figure 7), we became also interested in the potential of this enzyme to perform initial ZNA dependent DNA polymerase and ZNA dependent ZNA polymerase activities. This was tested using **ON11** (DNA-(*S*)-ZNA-DNA) (Table S2) as template and (*S*)-PMGApp as substrate. Notably, Klenow (exo$^-$) DNA polymerase could extend the primer by one (*S*)-PMGA residue within 1 h, while no (*R*)-PMGApp was incorporated (Figure 8), keeping the reaction stereoselective. Likewise, dATP is a successful substrate in this reaction.

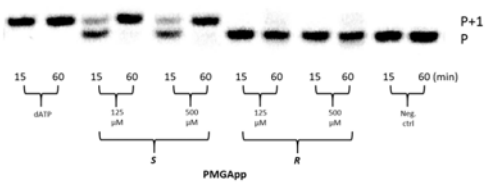


Figure 8. Profiles of incorporation of dATP and (*S*)/(*R*)-PMGApp by the Klenow (*exo*$^-$) fragment DNA polymerase (0.05 U.μL^{-1}) with Primer 2 (Table S10) and **ON11** as template.

In order to implement ZNA as genetic polymer *in vivo*, it is crucial that the constituting PMGNpp nucleotides are taken up by bacteria and utilized as substrates by evolved polymerases. The uptake and initial polymerization of PMGNpp *in vivo* was tested using an *E. coli* strain that has been previously re-engineered with an algal plastid transporter expressed from a synthetic gene inserted in the chromosome.[31] Specifically, such strain named XE858 and expressing the *ntt*2 gene encoding the *Thalassiosira pseudonana* transporter *Tp*NTT2 was proven capable of importing canonical dNTPs as well as modified nucleoside triphosphates. Both (*S*) and (*R*) diphosphates of acyclic nucleoside phosphonates featuring cytosine as nucleobase **7a** and **7b** were evaluated for uptake (transport) by testing their toxic effects on the XE858 strain growth on plates. In this test system, both compounds could inhibit cell growth at a 100 mM concentration in the well, indicating that both nucleotides are taken up by the XE858 *E. coli* strain and are substrates of an intracellular DNA polymerase (Figure 9). However, the (*S*)-PMGCpp was more toxic than (*R*)-PMGCpp leading to a comparatively larger halo. Due to its higher toxicity, and thus better uptake and incorporation into DNA, the (*S*) isomer seems to be a better candidate than the (*R*) isomer as a genetic polymer *in vivo*. No toxicity effect of compounds **7a** and **7b** was observed on the XE763 strain devoid of transporter (picture not shown). Similar results were obtained with enantiomeric purine analogues (*S*)-PMGApp **6a** and (*R*)-PMGApp **6b**, as detailed in the Supporting Information (Figure S8).

To further examine which polymerase could be involved in the incorporation of PMGCpp into *E. coli* DNA, DNA polymerase I and the Klenow fragment of DNA polymerase I were distinctly overexpressed in the strain bearing the algal plastid transporter and the corresponding variations of the toxicity effect due to the *in vivo* uptake of compounds **6a,b** and **7a,b** were thus determined. The choice for DNA polymerase Pol I was based on the initial *in vitro* incorporation study (Figures 7 and S5), where Pol I showed better incorporation of the modified nucleotides into DNA than Eco Pol III, although we realize that also other Eco polymerases could be involved in the process. Overexpression of Pol I in the XE858 strain conducted to an increased toxicity of the four PMGN diphosphates (Supplementary Information, Figures S9 and S10), as shown by an increase in the halo size or intensity.

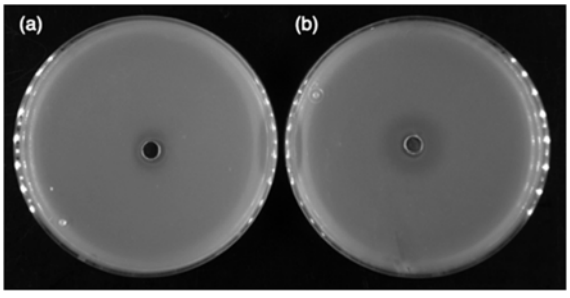


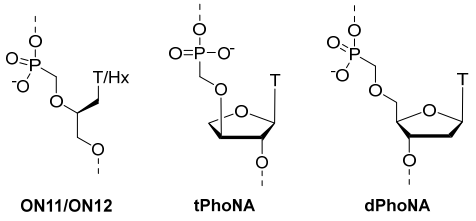
Figure 9. Toxic effect due to PMGCpp import visualized by an inhibition halo in the center of the plate. Response of *E. coli* strain XE858 expressing the ntt2 gene to (a) (*R*)-PMGCpp **7a** (halo 11 mm) and (b) (*S*)-PMGCpp **7b** (halo 18 mm).

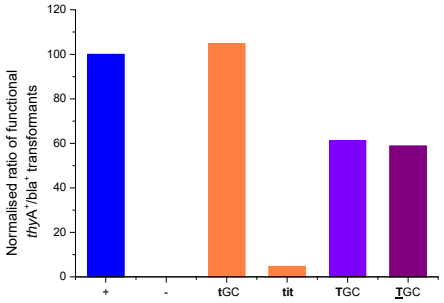
In recent studies, we investigated the ability of XNA/DNA mosaics to serve as templates for DNA synthesis *in vivo* using a nutritional selection scheme.[15] It was found that the genetic information carried by some modified six- and five-membered ring backbones [i.e., hexitol nucleic acid (HNA), cyclohexenyl nucleic acid (CeNA), and arabinofuranosyl nucleic acid (AraNA), and phosphomethylthreosyl nucleic acid

(tPhoNA)] could be tolerated by the replication machinery of *E. coli*, however no information transfer occurred between xenobiotic GNA and natural DNA.

Thus, (*S*)-ZNA/DNA oligonucleotide chimeras **ON11-12** (Table S2) were also synthesized (for synthetic procedure and compound characterization see Supporting Information, including Scheme S6) and tested using the same genetic selection scheme (Figure S12). The ratio between the bacterial colony numbers in + dt and - dt media indicates the success of the DNA-(*S*)-ZNA-DNA mosaics to serve as templates for DNA propagation *in vivo* (Table 2). As expected, the ligation of the Cys146 deleted oligonucleotide did not yield any prototrophic transformant. The same result was obtained when no oligonucleotide was added to the gapped vector. The replacement of a 2'-deoxyribose-5'-phosphate by one acyclic phosphonate as in **ON11** yielded a similar ratio of prototrophic transformants as the positive control (Table 2, Table S2). Further comparison with the only previous example of phosphono XNA, i.e., tPhoNA, as well as 5′-*O*-phosphonomethyl-deoxyribosyl nucleic acid (dPhoNA), which is structurally closer to a natural backbone, showed a more productive transliteration for **ON11** by *E. coli* polymerases at the level of a single modification. A 21-fold drop in the yield of prototrophic transformants resulted upon extension of the (*S*)-ZNA stretch from one to three oligonucleotides (**ON12**) (Table 2, Table S2). Thus, while a single acyclic phosphonate modification is well tolerated, the addition of successive acyclic phosphonate nucleotides reduces significantly the DNA propagation *in vivo*.

Table 2. Templating by mosaic DNA-(*S*)-ZNA-DNA oligonucleotides for restoring an active thymidylate synthase gene *in vivo* as compared with known phosphonate XNAs.





ON Sequences[(a)]								
	Leu[143]	Ala[144]	Pro[145]	Cys[146]	Hys[147]	Ala[148]	No. XNA residues	Ratio *thyA*$^+$/*bla*$^+$[(b)]
WT	CTA	GCG	CCG	TGC	CAT	GCA	0	0.085
ΔCys146	CTA	GCG	CCG	ΔΔΔ	CAT	GCA	0	0
ON11	CTA	GCG	CCG	**t**GC	CAT	GCA	1	0.089
ON12	CTA	GCG	CCG	**tit**	CAT	GCA	3	0.004
tPhoNA[20]	CTA	GCG	CCG	**T**GC	CAT	GCA	1	0.052
dPhoNA[20]	CTA	GCG	CCG	**T**GC	CAT	GCA	1	0.05

[(a)]Within each sequence, bases of acyclic, threosyl, and deoxyribosyl phosphonate nucleotides are indicated by lowercase bold, uppercase bold, and uppercase bold underlined letters, respectively; [(b)]*thyA*$^+$/*bla*$^+$ refers to the ratio of thymidine prototrophic colonies against the total number of transformant colonies. It should be noted that data for different oligonucleotides were obtained by performing independent experiments.

Since it was shown that the replacement of guanine in the TGC/T codon (encoding for the essential Cys146) by an hypoxanthine does not interfere with the correct copying of DNA,[15] the decrease of active *thyA* genes among the clones transformed owing to the extension of the acyclic phosphonate nucleotide stretch is most probably a result of an inefficient copying by *E. coli* polymerases rather than an erroneous addition. In conclusion, up to three contiguous acyclic phosphonate nucleotides were found to serve as a short template for *E. coli* replication enzymes, which represent the first demonstration that an acyclic oligonucleotide may be recognized by the *E. coli* replication system and converted into DNA *in vivo*. However, it should be noted that the biological read-out is biased by the specific codon and does not provide further information about potential misreading and base-base recognition.

DISCUSSION

Despite being a familiar concept in the area of nucleoside medicinal chemistry,[32] the use of the enzymatically stable phosphonate motif to replace the labile P-O bond for the selection of backbone modified nucleic acids as potential XNAs *in vitro* and *in vivo* has remained uninvestigated until recently. In an isolated example, threose phosphono nucleic acid (tPhoNA) was shown to poorly hybridize with canonical nucleic acids, while a single tPhoNA modified codon could be transliterated to DNA *in vivo*.[20] The viability of tPhoNA to serve as genetic polymer *in vitro* was demonstrated by evolving dedicated polymerases to copy tPhoNA into DNA and back into tPhoNA. However, considering the synthetic difficulty we encountered in producing large amounts of the corresponding monomers, it was envisaged that an acyclic phosphono backbone, such as that shown in Figure 1, could equally hold the potential to address the need for XNAs that are stable against enzymatic hydrolysis, while being synthetically more accessible. Moreover, dealing with an increasingly altered backbone chemistry could be expected to be accompanied by a higher orthogonality gain (i.e., lower hybridization with natural nucleic acids). Previous studies showed that the acyclic backbone of GNA could not be recognized and used as template for DNA synthesis by the replicative enzymes of *E. coli*.[15] However, the transliteration efficiency by cellular polymerases for the proposed phosphonate analogue could not be predicted based on this precedent. Phosphonate nucleosides in fact require only two phosphorylation steps, instead of three, for their metabolic activation.

The solid-phase chemical synthesis of phosphonate-linked oligonucleotides is not as straightforward as that of oligonucleotides with a phosphodiester backbone; for instance, the standard phosphonamidite approach was not practicable in this case. However, after extensive experimentation, a satisfactory protocol was developed by further adapting a modified *H*-phosphonate literature protocol,[33] which enabled the assembly of phosphonomethylglycerol units with either a (*S*)- or (*R*)-configuration into enantiomeric oligomers. On the other hand, the DNA fragments of these chimeric oligonucleotides were synthesized by using the classical phosphoramidite method. According to thermal stability studies of homo- and hetero duplexes, the insertion of a single acyclic phosphonate nucleotide in a DNA sequence produced a significant reduction in the stability of the corresponding duplexes, regardless of the chirality of the modified nucleotide. In contrast, fully modified homoduplexes of mixed sequences containing A, T, C, and G proved to be more stable than dsDNA. It is interesting to note that oligonucleotides built from (*S*)-monomers [i.e., (*S*)-ZNA] hybridizes weakly with DNA, while their (*R*)-counterparts [(*R*)-ZNA] completely fails to hybridize with DNA.

A closer examination of the helicity of (*S*)-dsZNA, (*R*)-dsZNA, and (*S*)-ZNA:DNA heteroduplexes by CD spectroscopy and molecular modeling suggested that (*S*)-dsZNA formed a left-handed helix, while (*R*)-dsZNA gave rise to a right-handed helix. The DNA coil of the (*S*)-ZNA:DNA helix occurred to be slightly left-handed with large deviations of the α and ζ backbone angles, when compared with a regular dsDNA helix. This is a remarkable observation, especially considering that (*S*)-ZNA with a tendency to form left-handed helix hybridizes with DNA, while the right-handed helix of (*R*)-ZNA did not lead to the formation of DNA hybrids. However, this observation is preliminary and not conclusive. It has been observed previously that (*S*)-GNA which shows a similar CD spectra as (*S*)-ZNA, adopts a right-handed helix upon crystallization. The fact that (*S*)-ZNA preserves hybridization properties with DNA allows to use present biotechnological tools for its further development.

Finally, to explore the possibility of using (*S*)-ZNA as information system for subsequent evolution processes, it is of crucial interest to determine its *in vivo* recognition properties. PMGN diphosphates of both enantiomeric series were shown to serve as substrates for Klenow (*exo*-) DNA polymerase with nucleobase selectivity. Nonetheless, the fact that the (*S*)-monomers can be recognized by this polymerase and incorporated into DNA regardless of the inherent predisposition of (*S*)-ZNA to potentially adopt a left-handed form, it is a true evidence to the remarkable properties of such XNA. Encouraging for potential *in vivo* replication, is the observation that Klenow (exo$^-$) DNA polymerase shows initial ZNA dependent ZNA polymerase activity. Additionally, (*S*)-PMGCpp and (*S*)-PMGApp were accepted by an algal plastid transporter expressed in *E. coli*, as demonstrated by cell growth inhibition following their uptake. Overexpression of polymerase I of *E. coli* in the strain encoding the transporter led to increased toxicity. Notably, it was demonstrated that (*S*)-ZNA could be transliterated into DNA in *E. coli* at the level of one codon, which represents a first example of the recognition of an acyclic oligonucleotide by a cellular polymerase *in vivo*.

CONCLUSION

In summary, the chemical synthesis of a xenobiotic nucleic acid with an acyclic pseudosugar-phosphonate backbone has been achieved starting from readily available and inexpensive chiral phosphonomethylglycerol nucleosides. (*S*)-ZNA exceptionally exemplifies a chemically orthogonal information system owing to its strong self-hybridization properties along with a weak interaction with complementary single-stranded DNA. Diphosphates of (*S*)-ZNA constituting monomers were identified as substrates by *E. coli* polymerase I as well as an algal nucleotide transporter. The transliteration of (*S*)-ZNA in *E. coli* was demonstrated at the level of a single codon, however this result successfully highlights the templating potential of acyclic oligonucleotides for subsequent *in vivo* evolution.

ASSOCIATED CONTENT

The Supporting Information is available free of charge on the ACS Publications website.

Detailed synthesis procedures, characterization information, NMR spectra, supplementary methods, figures, and references (PDF).

AUTHOR INFORMATION

Corresponding Author

* piet.herdewijn@kuleuven.be

Present Addresses

†Faten Jaziri, DNA Script, 29 Rue du Faubourg Saint Jacques, 75014, Paris, France.

Author Contributions

All authors have given approval to the final version of the manuscript.

ACKNOWLEDGMENT

M.L. acknowledges the China Scholarship Council (CSC) for funding (grant 201306220071). 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. ERC2012-ADG 20120216/320683). We also 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 express our gratitude to Luc Baudemprez, Lia Margamuljana, and Hoai Nguyen for excellent technical assistance.

REFERENCES

1. Pinheiro, V. B.; Holliger, P., The XNA World: Progress Towards Replication and Evolution of Synthetic Genetic Polymers. *Curr. Opin. Chem. Biol.* **2012**, *16*, 245-252.
2. Chen, T.; Hongdilokkul, N.; Liu, Z. X.; Thirunavukarasu, D.; Romesberg, F. E., The Expanding World of DNA and RNA. *Curr. Opin. Chem. Biol.* **2016**, *34*, 80-87.
3. Herdewijn, P.; Marlière, P., Toward Safe Genetically Modified Organisms through the Chemical Diversification of Nucleic Acids. *Chem. Biodiversity* **2009**, *6*, 791-808.
4. Malyshev, D. A.; Dhami, K.; Lavergne, T.; Chen, T.; Dai, N.; Foster, J. M.; Corrêa, I. R.; Romesberg, F. E., A Semi-Synthetic Organism With an Expanded Genetic Alphabet. *Nature* **2014**, *509*, 385-388.
5. 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., A Semi-Synthetic Organism That Stores and Retrieves Increased Genetic Information. *Nature* **2017**, *551*, 644-647.
6. Marlière, P.; Patrouix, J.; Döring, V.; Herdewijn, P.; Tricot, S.; Cruveiller, S.; Bouzon, M.; Mutzel, R., Chemical Evolution of a Bacterium's Genome. *Angew. Chem., Int. Ed.* **2011**, *50*, 7109-7114.
7. Mehta, A. P.; Li, H.; Reed, S. A.; Supekova, L.; Javahishvili, T.; Schultz, P. G., Replacement of 2'-Deoxycytidine by 2'-Deoxycytidine Analogues in the E. coli Genome. *J. Am. Chem. Soc.* **2016**, *138*, 14230-14233.
8. Mehta, A. P.; Li, H.; Reed, S. A.; Supekova, L.; Javahishvili, T.; Schultz, P. G., Replacement of Thymidine by a Modified Base in the Escherichia coli Genome. *J. Am. Chem. Soc.* **2016**, *138*, 7272-7275.
9. Eremeeva, E.; Abramov, M.; Margamuljana, L.; Rozenski, J.; Pezo, V.; Marliere, P.; Herdewijn, P., Chemical Morphing of DNA Containing Four Noncanonical Bases. *Angew. Chem., Int. Ed.* **2016**, *55*, 7515-7519.
10. Krueger, A. T.; Peterson, L. W.; Chelliserry, J.; Kleinbaum, D. J.; Kool, E. T., Encoding Phenotype in Bacteria with an Alternative Genetic Set. *J. Am. Chem. Soc.* **2011**, *133*, 18447-18451.
11. El-Sagheer, A. H.; Sanzone, A. P.; Gao, R.; Tavassoli, A.; Brown, T., Biocompatible Artificial DNA Linker that is Read through by DNA Polymerases and is Functional in Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 11338-11343.
12. Sanzone, A. P.; El-Sagheer, A. H.; Brown, T.; Tavassoli, A., Assessing the Biocompatibility of Click-Linked DNA in Escherichia coli. *Nucleic Acids Res.* **2012**, *40*, 10567-10575.
13. Birts, C. N.; Sanzone, A. P.; El-Sagheer, A. H.; Blaydes, J. P.; Brown, T.; Tavassoli, A., Transcription of Click-Linked DNA in Human Cells. *Angew. Chem., Int. Ed.* **2014**, *53*, 2362-2365.
14. 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., Synthetic Genetic Polymers Capable of Heredity and Evolution. *Science* **2012**, *336*, 341-344.
15. Pezo, V.; Liu, F. W.; Abramov, M.; Froeyen, M.; Herdewijn, P.; Marliere, P., Binary Genetic Cassettes for Selecting XNA-Templated DNA Synthesis In Vivo. *Angew. Chem., Int. Ed.* **2013**, *52*, 8139-8143.
16. Inoue, N.; Shionoya, A.; Minakawa, N.; Kawakami, A.; Ogawa, N.; Matsuda, A., Amplification of 4'-ThioDNA in the Presence of 4'-Thio-dTTP and 4'-Thio-dCTP, and 4'-ThioDNA-Directed Transcription In Vitro And In Mammalian Cells. *J. Am. Chem. Soc.* **2007**, *129*, 15424-15425.
17. Maruyama, H.; Furukawa, K.; Kamiya, H.; Minakawa, N.; Matsuda, A., Transcription of 4'-ThioDNA Templates to Natural RNA in Vitro and in Mammalian Cells. *Chem. Commun.* **2015**, *51*, 7887-7890.
18. Holy, A., Antiviral Acyclic Nucleoside Phosphonates Structure Activity Studies. *Antiviral Res.* **2006**, *71*, 248-253.
19. Zhang, L. L.; Peritz, A.; Meggers, E., A Simple Glycol Nucleic Acid. *J. Am. Chem. Soc.* **2005**, *127*, 4174-4175.
20. Liu, C.; Cozens, C.; Jaziri, F.; Rozenski, J.; Marechal, A.; Dumbre, S.; Pezo, V.; Marliere, P.; Pinheiro, V. B.; Groaz, E.; Herdewijn, P., Phosphonomethyl Oligonucleotides as Backbone-Modified Artificial Genetic Polymers. *J. Am. Chem. Soc.* **2018**, *140*, 6690-6699.
21. An, H. Y.; Wang, T. M.; Maier, M. A.; Manoharan, M.; Ross, B. S.; Cook, P. D., Synthesis of Novel 3'-C-Methylene Thymidine and 5-Methyluridine/Cytidine H-Phosphonates and Phosphonamidites for New Backbone Modification of Oligonucleotides. *J. Org. Chem.* **2001**, *66*, 2789-2801.
22. Schneider, K. C.; Benner, S. A., Oligonucleotides Containing Flexible Nucleoside Analogs. *J. Am. Chem. Soc.* **1990**, *112*, 453-455.
23. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E., UCSF Chimera-A Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004**, *25*, 1605-1612.
24. Zgarbova, M.; Otyepka, M.; Sponer, J.; Lankas, F.; Jurecka, P., Base Pair Fraying in Molecular Dynamics Simulations of DNA and RNA. *J. Chem. Theory Comput.* **2014**, *10*, 3177-3189.
25. Schlegel, M. K.; Essen, L.-O.; Meggers, E., Duplex Structure of a Minimal Nucleic Acid. *J. Am. Chem. Soc.* **2008**, *130*, 8158-8159.
26. Schlegel, M. K.; Essen, L.-O.; Meggers, E., Atomic Resolution Duplex Structure of the Simplified Nucleic Acid GNA. *Chem. Commun.* **2010**, *46*, 1094-1096.
27. Florián, J.; Štrajbl, M.; Warshel, A., Conformational Flexibility of Phosphate, Phosphonate, and Phosphorothioate Methyl Esters in Aqueous Solution. *J. Am. Chem. Soc.* **1998**, *120*, 7959-7966.
28. Fijalkowska, I. J.; Schaaper, R. M.; Jonczyk, P., DNA Replication Fidelity in Escherichia coli: A Multi-DNA Polymerase Affair. *FEMS Microbiol. Rev.* **2012**, *36*, 1105-1121.
29. Horhota, A. T.; Szostak, J. W.; McLaughlin, L. W., Glycerol Nucleoside Triphosphates: Synthesis and Polymerase Substrate Activities. *Org. Lett.* **2006**, *8*, 5345-5347.
30. Chen, J. J.; Tsai, C. H.; Cai, X.; Horhota, A. T.; McLaughlin, L. W.; Szostak, J. W., Enzymatic Primer-Extension With Glycerol-Nucleoside Triphosphates on DNA Templates. *PloS one* **2009**, *4*, e4949.
31. Pezo, V.; Hassan, C.; Louis, D.; Sargueil, B.; Herdewijn, P.; Marlière, P., Metabolic Recruitment and Directed Evolution of Nucleoside Triphosphate Uptake in Escherichia coli. *ACS Synth. Biol.* **2018**, *7*, 1565-1572.
32. De Clercq, E.; Holy, A., Acyclic Nucleoside Phosphonates: A Key Class of Antiviral Drugs. *Nat. Rev. Drug Discovery* **2005**, *4*, 928-940.
33. Kostov, O.; Pav, O.; Budesinsky, M.; Liboska, R.; Simak, O.; Petrova, M.; Novak, P.; Rosenberg, I., 4-Toluenesulfonyloxymethyl-(H)-phosphinate: A Reagent for the Introduction of O-and S-Methyl-(H)-phosphinate Moieties. *Org. Lett.* **2016**, *18*, 2704-2707.

For the Table of Contents Only