Organic &
Biomolecular
Chemistry

# Achiral, acyclic nucleic acids; Synthesis and biophysical studies of a possible prebiotic polymer

SCHOLARONE™
Manuscripts

## ARTICLE

# Achiral, acyclic nucleic acids; Synthesis and biophysical studies of a possible prebiotic polymer

P. Srivastava,[a] R. Abou El Asrar,[a] C. Knies,[b] M. Abramov,[a] M. Froeyen,[a] J. Rozenski,[a] H. Rosemeyer,[b] and P. Herdewijn[a*]

The search for prebiotic, nucleic acid precursors is, at its best, a speculative undertaking. Given the complex structure of RNA, it is not very likely that RNA was the first information system in the universe and thus finding possible precursor/s i.e. pre-RNA remains an open challenge. We, in this paper, have tried to construct nucleic acid polymers with simple acyclic, achiral backbone. Such a linear, achiral backbone may have been formed from simple monomers that may have existed in the ''prebiotic soup''. We have shown that such polymers are capable of identifying the complementary "other self" and thus forming a potential system for information storage and transmission. This study thus involves investigation of nucleic acid analogues with modified backbone that are likely to have formed in the prebiotic setting.

## Introduction

"*Supporters of the hypothesis that RNA was the first genetic material must explain where the nucleotides came from and how they were self-organized. Those who believe in a simple precursor have a difficult task of identifying such a precursor, but they hope that explaining monomer synthesis will then be simpler, L. Orgel*"[1]

The space between the classic Urey-Miller experiment on the origin of biomolecules[2,3] and the "RNA-world"[4] leaves much room for possible permutation combinations in the search for a pre-RNA ancestor/ancestors.[5] With the "RNA world" theory rather firmly established, the puzzle now is the origin of the RNA world.[6-8] What was/were the RNA precursor/s[9] that underwent chemical evolution and the "fittest" of them came to dominate the biotic landscape?[10] This would mean a longer list of "extinct" molecular species. A variety of organic moieties have been presented as RNA precursor/s i.e. the Pre-RNA,[11-15] one of them being PNA,[16] in the category of non-sugar nucleobase conjugates (Figure 1B). It has also been proposed that a pre-RNA may have completely lacked sugar/s[17] and even phosphates[18] and thus a linear, achiral polymer conjugated to
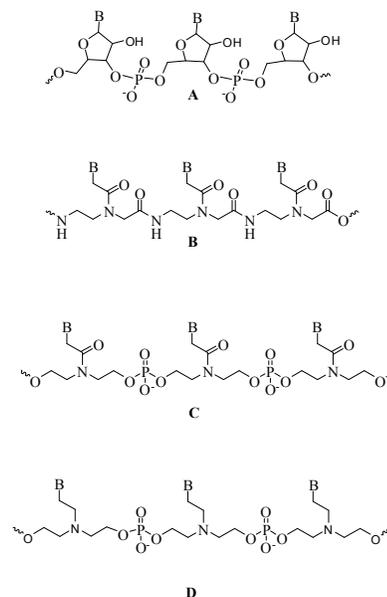


**Figure 1:** Representative examples of nucleic acid polymers; (**A**) RNA (**B**) PNA, (**C**) DEANA, (**D**) TEANA
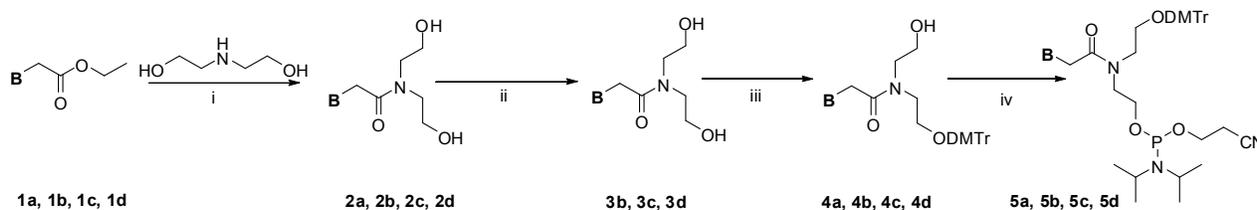
Scheme 1: Reagents and conditions: (i) 2,2'-Iminodiethanol, Ethanol, reflux, overnight; (ii) TMS-Cl, Benzoyl chloride/ Isobutryl chloride, dry pyridine, 0° C to r.t., 3 hr; (iii) DMTr-Cl, dry pyridine or dry DMF and Et₃N, 0° C to r.t., overnight; (iv) 2-Cyanoethyl N,N-diisopropylchlorophosphoramidite, DIPEA, dry DCM/THF, 0° C to r.t. 30 min.

B= Thymine (**1a-5a**), Cytosine (**1b, 2b**), Adenine (**1c, 2c**) and Guanine (**1d, 2d**), or C^Bz (**3b-5b**), A^Bz (**3c-5c**) and G^ibu (**3d-5d**) respectively.

nitrogen heterocycles might represent a more realistic prebiotic (pre-RNA) organic molecule capable of assembling and recognizing a distinct "partner".[19-22] Apart from their therapeutic potential[23,24] acyclic nucleic acid analogues have been proposed as candidates for the prebiotic, pre-RNA world.[25-30]

Herein, we report synthesis and biophysical evaluation of an achiral, acyclic nucleic acid analogue, named DEANA *i.e.* DiethanolAmideNucleicAcid having canonical nucleobases as recognition element with a flexible phosphodiester backbone (Figure 1C). In this paper, we will also introduce another analogue that we have named TEANA i.e. TriethanolAminoNucleicAcid having phosphorylated triethanolamine as the repeating unit engrafted with nucleobase (Figure 1D). The synthesis of TEANA was carried out on account of extensive degradation of DEANA oligomer during deprotection of the synthesized oligos. We are introducing early studies on TEANA with adenine, and full studies shall be reported in the future. Structurally similar analogues[31,32] e.g. PPNA[33] and PHONA[34] have been synthesized and the lower pairing affinity of such acyclic phosphonate oligomers has been reported.

The backbone of DEANA is composed of diethanolamine conjugated to the recognition element via amide bond formation between N-alkylated nucleobase acetic acids and the secondary amine. The acetic acid analogues of nucleobases may have been present in the primordial soup[35,36] and the reaction of ammonia with oxirane[37] would easily result in the formation of diethanolamine and triethanolamine in the same soup. We have thus attempted to construct oligomeric conjugates formed by possible prebiotic molecules i.e. nucleobase acetic acids and diethanolamine and, through the study of biophysical properties of such "prebiotic" polymers tried to expand the repertoire of potential pre-RNA constructs.

## Results and discussion

Our scheme for the synthesis of DEANA (Scheme 1) starts with the regioselective alkylation of nucleobases, i.e. thymine, adenine, cytosine and guanine by ethyl bromoacetate in DMF

using the established synthetic procedure.[38] The appropriately alkylated ethyl esters of thymine-1-yl acetic acid, cytosine-1-yl acetic acid, adenine-9-yl acetic acid, and guanine-9-yl acetic acid were refluxed with 2,2'-iminodiethanol overnight in ethanol as reported for synthesis of N,N-bis-(2-hydroxyethyl)-2-(thymin-1-yl)acetamide[39] to obtain N,N-bis-(2-hydroxy-ethyl)-2-(thymin-1-yl)acetamide (**2a**), N,N-bis-(2-hydroxy-ethyl)-2-(cytosine-1-yl)acetamide (**2b**), N,N-bis-(2-hydroxyethyl)-2-(adenine-9-yl)acetamide (**2c**), and N,N-bis-(2-hydroxyethyl)-2-(guanine-1-yl)acetamide (**2d**) respectively, as white solids in nearly quantitative yields. The exocyclic amino functions of the nucleobases were protected as benzoyl and isobutanoic amides for adenine/cytosine and guanine bases, thus providing **3b**, **3c** and **3d** respectively . One of the hydroxyls was protected as dimethoxytrityl ether (4a, 4b, 4c, 4d) and the remaining hydroxyl converted into phosphoramidite with 2-cyanoethyl N,N-diisopropyl-chlorophosphoramidite, thus providing the monomer phosphoramidites (**5a**, **5b**, **5c**, **5d**) for further solid phase oligomer synthesis using the standard, automated solid support synthetic protocol on Expedite® RNA/DNA synthesizer. 5-(Ethylthio)tetrazole was used as the activating agent with 10 min coupling time for all DEANA phosphoramidites. This was followed by de-protection from solid support and base deprotection in ammonia at 55° C followed by purification to obtain fully modified (DEANA) or chimeric (DNA-DEANA) oligonucleotides.

Synthesis of the triethanolamine nucleoside (Scheme 2) starts with regioselective alkylation of adenine with 1,2-dibromoethane as reported.[40] The 9-(2-bromoethyl)-adenine (**6a**) was then refluxed in dioxane with silyl protected 2, 2'-iminodiethanol and Cesium carbonate to obtain **6b**. The exocyclic amino group of the compound **6b** was then protected as bis-benzamide using benzoyl chloride in pyridine, thus providing **6c**.

Triethylamine trihydrofluoride was then used to deprotect the silyl ethers to obtain bis-hydroxyl product **6d**. One of the hydroxyls was next converted into a dimethoxytrityl ether (**6e**) and the other was converted into a phosphoramidite with 2-cyanoethyl N,N-diisopropylchlorophosphoramidite, thus providing the adenine-containing TEANA phosphoramidite monomer**6f** for further solid phase oligomer synthesis using the
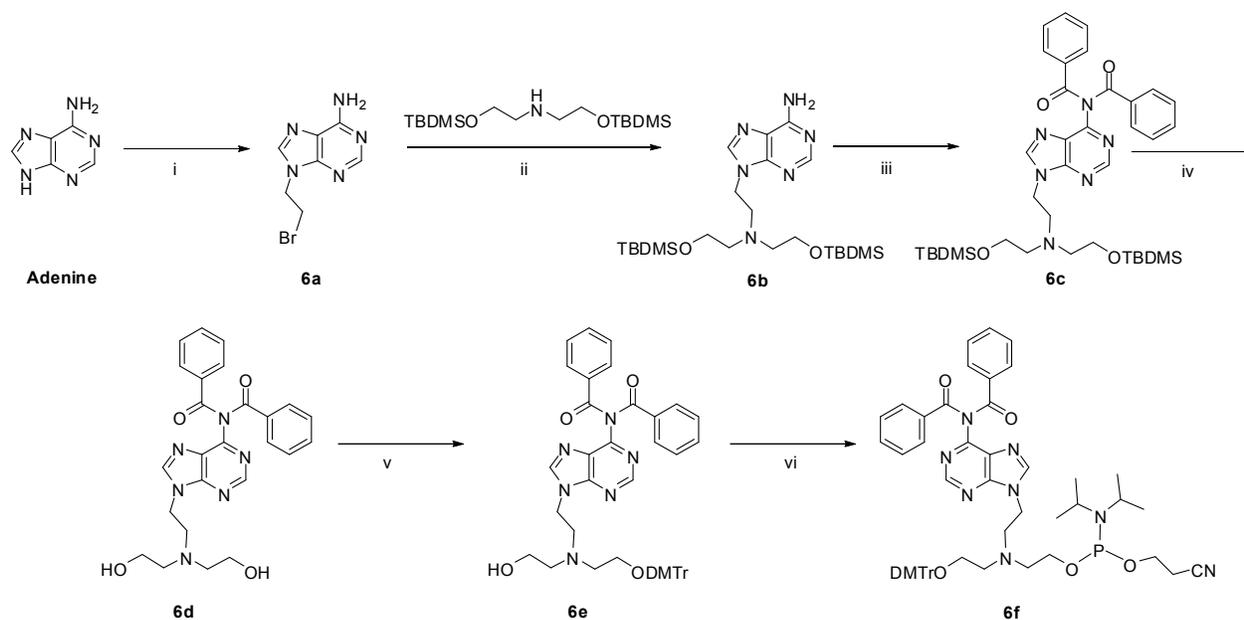
**Scheme 2 :**Reagents and conditions: (i) 1,2-dibromoethane, $K_2CO_3$, dry DMF; (ii) $CsCO_3$, dry dioxane, reflux, 12 h; (iii) Benzoyl chloride, dry pyridine, 0° C to r.t. overnight; (iv) triethylamine trihydrofluoride, dry THF, 0°C to r.t. 6-8 hr; (v) DMTr-Cl, dry pyridine, 0°C to r.t., overnight; (vi) 2-Cyanoethyl N,N-diisopropylchlorophosphoramidite, DIPEA,dry THF, 0° C to r.t. 30 min.

standard, automated solid support, synthetic protocol on Expedite® RNA/DNA synthesizer. 5-(Ethylthio)tetrazole was used as the activating agent with 10 min. of coupling time for TEANA phosphoramidite followed by deprotection andpurification to obtain the singly modified, TEANA/DNA mixed oligonucleotide.

**Thermal melting studies:**

The base pairing properties of DEANA and TEANA modified oligomer sequences were investigated by temperature dependent UV absorption at 260 nm wavelength. In our initial attempt, the fully modified 15-mer thymine containing DEANA sequence (ON-1) showed no duplex melting transition with a 15-mer-dA sequence. To find out if the analogues show self-pairing, a fully modified the 15-mer sequence was synthesized with the phosphoramidite monomer **5c** (Scheme 1).

Table 1[a]

| Oligo. | Sequence | Tm in °C | Mass-found/calculated* |
|---|---|---|---|
| ON-1 | **(t)**$_{15}$: d(A)$_{15}$ | N.A | 4936.3/4936.5 |
| ON-2 | d(CGCGAATTCGCG)$_2$ | 51.5 | - |
| ON-3 | d(CG**cgaattcg**CG)$_2$ | 27.5 | 3878.8/3878.7 |
| ON-4 | [CH$_3$O(CH$_2$CH$_2$O)$_2$**gcgcgcgcg**C]$_2$ | 29.5 | 3487.1/3487.4 |

[a]Thermal melting studies of oligonucleotides containing natural/modified nucleic acid analogues (DEANA). Modifications are represented by lower case bold letters (**a, t, g, c**) in the oligonucleotide sequence. Change in absorbance measured at 260 nm in NaCl (0.1M) buffer with KH$_2$PO$_4$ (20 mM, pH 7.5) and EDTA (0.1 mM) at the concentration of 4 μM for each strand and 8 μM for the self-complementary Dickerson Drew dodecamer (ON-3) and ON-4. *Mass of modified sequence

However, the fully modified adenine containing DEANA was hydrolyzed during the base deprotection step, probably *via* trans-esterification, analogous to the mechanism of RNA degradation. Previously,[30] we had observed similar problems while trying to synthesize other acyclic nucleic acids. In the case of DEANA an intranucleotide transesterification reaction in addition to the cleavage of the amide bond between nucleobase acetic acid and diethanolamine could have occurred. Taking clue from other acyclic systems, we next synthesized a modified Dickerson-Drew dodecamer[41] with two flanking deoxyribo nucleotides at the 3' and 5' ends, thus avoiding decomposition of the DEANA containing oligonucleotide by trans-esterification in the basic medium. This self-complementary sequence was chosen since it has been extensively studied[42] for gaining insights into the structure of nucleic acids.[43-47] The temperature-dependent UV absorption study of the chimeric Dickerson Drew dodecamer sequence at 260 nm exhibited a sigmoidal melting curve (Figures SI-3 and SI-4, in supporting information) at 0.1M NaCl concentration thus providing a $T_m$ of 27.5 °C (Table 1) corresponding to an approximately 24 °C drop in Tm as compared to the all DNA-dodecamer (ON-2). This accounts to an approximately 1.5 °C drop in $T_m$ per modification at the same salt and oligo concentration. Realizing that end capping was necessary we next synthesized another self-complementary oligomer (ON-4) with diethyleneglycol monomethyl ether as 5' end cap and a natural dC as 3' end cap (also see the discussion on CD studies). The yields of oligonucleotide synthesis was better after deprotection as compared to the results obtained with fully modified oligomers, similar to ON-3. This self-complementary sequence (ON-4) exhibited a $T_m$ of 29.5° C with no hysteresis

between heating and annealing curves (Figure SI-5 in supporting information). To further understand the hybridization process a concentration dependent $T_m$ analysis was done for ON-4. The melting temperature for ON-4 was found to be constant from 2μM to 12 μM concentrations (Figure-SI-7 in supporting information) thus indicating a unimolecular process. Given the acyclic nature of the DEANA it is not unlikely that we have a hairpin at our hands (also see the molecular modeling section). Thus, we were able to show that DEANA is capable of self-pairing albeit with low thermal stability. The lability of the carboxamide bond guided us towards the synthesis of TEANA and thus a stable N-alkyl linker was substituted for a labile amide linker in DEANA. In a priliminary study, a single TEANA modification was introduced in a DNA sequence. The thermal melting study (Figure SI-6 in Supporting information) of this modified oligo with the complementary DNA sequence exhibited a drop of 7.0° C in $T_m$ as compared to the natural all DNA:DNA duplex of the same sequence (Table 2)

Table 2[a]

| Sequence-duplex | Tm in °C | Mass-found/calculated[*] |
|---|---|---|
| 5'-d(CTTCATTTTTTCTTC)-3' 3'-d(GAAGTAAAAAAGAAG)-5' | 47.0 °C | - |
| 5'-d(CTTC**a**TTTTTTCTTC)-3' 3'-d(GAAGTAAAAAAGAAG)-5' | 40.0 °C | 4462.9/4462.8 |

[a]Thermal melting studies of oligonucleotides containing natural/ modified nucleic acid analogues. Modification is represented by lower case bold letter (**a**) in the oligonucleotide sequence. Change in absorbance measured at 260 nm in NaCl (0.1m) buffer with $KH_2PO_4$ (20 mM, pH 7.5) and EDTA (0.1 mM) at the concentration of 4 μM for each strand. *Mass of modified sequence

Since the bis-hydroxyethanol-phosphodiester backbone in DEANA has a higher degree of torsional flexibility, we had hoped that a stable duplex formation would be possible by flexing to maximize stacking interactions and hydrogen bonding. The decrease in stability when comparing DEANA to DNA can be explained by the entropic penalty[48,49] for the intra system base pairing. The thermal melting study of homoduplex with our acyclic analogue (DEANA) is in contrast with a previous study[50] wherein the homoduplexes of other acyclic analogues exhibited higher thermal stability. It is known that pairing (self or cross) of oligonucleotides is determined by the backbone inclination.[51] Hence, to gain insights into the structural features of DEANA oligonucleotides, molecular modeling studies were done.

**Molecular modeling studies:**

In order to gain insight into the DEANA duplex structure and to answer the question as to how a nucleic acid analogue with 8-bonds between the repeating nucleotides could be formed instead of 6 bonds in the natural DNA, we constructed a static model of a right-handed DEANA double stranded helix with

sequence CGTACG:CGTACG as described in the experimental section. It should be noted that the DEANA nucleotide has no 5'end different from 3'end, so duplexes cannot be labeled parallel or anti-parallel (Figure 2).

The static model shows that two strands with DEANA residues can be combined to form double stranded self-pairing duplexes with Watson Crick base pairing geometry (Figure 2A) which is also supported by the CD spectra of the dodecamer flanked by natural deoxy residues (see further in the CD studies). The modeled DEANA duplex adopts a structure very similar to the experimental PNA structure with 18 base pairs per turn forming an elongated helix.[52] The phosphates are expelled more outside of the helix than the phosphates in the classical DNA helix (Figure 2B). However, the pairing property of DEANA is markedly inferior to that of PNA, wherein the PNA duplexes give higher melting temperatures than DNA duplexes.[53] In addition to the previously mentioned entropic penalty, this may be explained by the lack of phosphate groups in PNA (no repulsion of the single strands in the duplex) and the inter-strand stacking due to the high inclination. The DEANA has charged phosphate groups, so the enthalpy contribution to the hybridization energy is expected to be lower due to the electrostatic repulsion. It must however be noted that the model was created by superposition of the Watson-Crick base pairs of DEANA onto the PNA bases and then minimizing the energy of the duplex, resulting in the creation of the phosphate bulges, and almost not changing the base positions. It must also be noted that Dickerson-Drew dodecamer is known to exist in duplex and hairpin forms, No attempt however was made to dissect the molecularity of hybridization for ON-3.

Table 3[a]

| | Handedness | Disp | Rise [Å] | Incl [°] | Tilt | Twist [°] | Bp/turn | P1-P1 dist | P1-P2 dist | P1-P3 dist |
|---|---|---|---|---|---|---|---|---|---|---|
| PNA | Right | 8.3 | 3.2 | 0.3 | 1.0 | 19.8 | 18 | - | - | - |
| DEANA | Right | 8.3 | 3.2 | 0.3 | 1.0 | 19.8 | 18 | 15.01 | 11.12 | 9.55 |
| B-DNA | Right | 0.0 | 3.4 | 2.4 | 0 | 36 | 10 | 18.14 | 16.69 | 14.93 |

[a]Helical parameters for PNA and B-DNA double helix.[52] Values for DEANA are copied from the PNA structure, except the P-P distances. P1-P1 is the distance in Å units between 2 phosphates of complementary nucleotides in the anti-parallel duplex, P1-P2 is the distance to the next P in the chain etc.

Table 3 shows the remarkable similarities in the helical parameters of PNA and a DEANA (ON-3) duplex, obtained from the superposition of the DEANA model onto the PNA structure (Figure 3) despite the fact that we have two extra atoms (or 2 extra bonds) in the backbone. Starting from the left handed PNA NMR structure (pdb structure 2K4G),[52] an attempt was made to create a left handed DEANA helix as well.
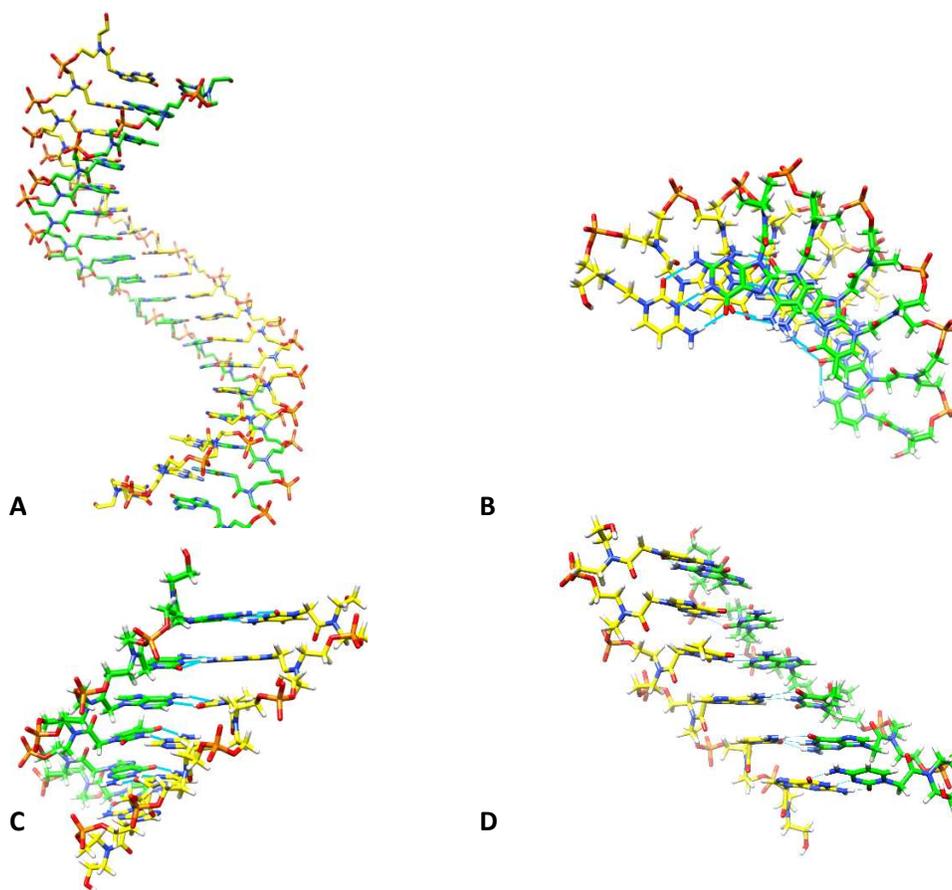
**Figure 2:** Static DEANA 18-mer duplex model structure. (**A**) Front view, (**B**) Top view shows the characteristic popping out of the phosphates, (**C**) View from minor groove, (**D**) Close-up showing hydrogen bonding and nucleobase stacking. Carbons of one strand are colored in yellow, in the other strand in green. Hydrogen bonds are shown as cyan lines.
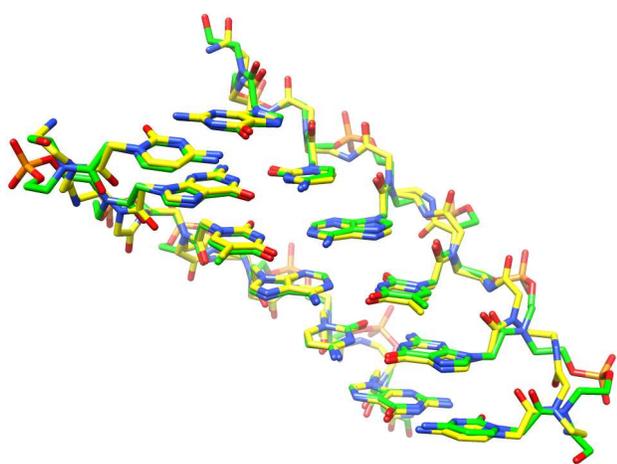


**Figure 3:** Superposition of the static DEANA model structure onto the PNA x-ray structure 1PUP, highlighting the popping out of the phosphate groups.

However, when optimizing the final structure in Amber, the amide connections from the main chain to the base showed high torsion energy with non-planer amide bonds. It must be noted that DEANA has no chiral center and hence left and right

handed helixes with the same geometry can be formed with equal probability and DEANA (nucleosides or oligonucleotide) on itself cannot contribute to symmetry breaking. Since the concentration dependent $T_m$ study for ON-4 indicated a unimolecular pairing process (Fig SI-7), and coupled with the absence of classical sigmoidal $T_m$ response, a hairpin model of 10mer **g.c** repeat sequence was constructed with DEANA using PNA double helix from the 3MBS pdb entry,[54] deleting all pairs except 3 GC pairs and adding a GCGC tetraloop (conformation not defined). The molecular mechanics energy of the loop part was optimized in Amber12. A DEANA hairpin was then constructed based on this PNA hairpin and energy minimized in Amber12, restraining the base atoms. With only three base pairs in the stem of a 10mer hairpin sequence (Figure SI-8 in supporting information), it is not surprising that we see only slight change in the absorbance during melting. A single strand melting of a non-hairpin structure, however, cannot be ruled out.

**Circular diachroism study:**

The self-complementary dodecamer sequence containing achiral acyclic DEANA (ON-3) synthesized with flanking natural deoxy residues (Table 1) was used to study if such

chimeric oligomer will adopt a discernible helical structure. Temperature dependent CD spectra were recorded in the neutral cacodylate buffer solution. As evident from the CD spectra, (Figure 4) the dodecamer adopts a right handed helix with maxima around 280 nm and minima around 250 nm. This signature ellipticity is analogous to that shown by the B-type DNA duplex. The $T_m$ of the dodecamer, as determined by the UV melting study, was also confirmed with this temperature dependant CD measurement.
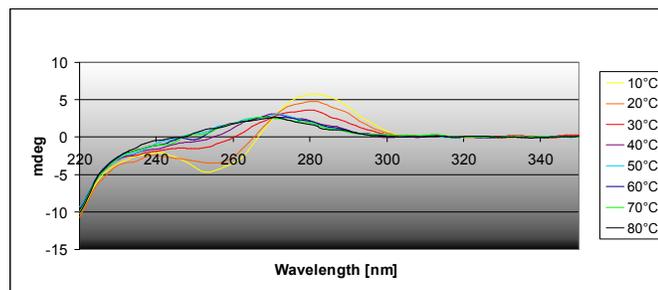


**Figure 4:** Temperature dependent CD spectra of dodecamer duplex sequence ON-3 recorded in cacodylate buffer.

An achiral, acyclic DEANA polymer was not expected to show this CD spectra and the flanking chiral deoxy residues have steered the acyclic DEANA to this structure in a well-known phenomenon of chiral induction[55] also observed in case of PNA wherein a terminal amino acid residue or deoxynucleotides residue/s drives the polymer to adopt right handed or left handed form.[56-60]Since we had in our hand ON-4 with only a single natural residue at 3' end, a CD study was warranted to show that the signature CD shown by the ON-3 was indeed a result of chiral induction. The ON-4 did not exhibit any ellipticity in the temperature dependent CD study (Figure 5).
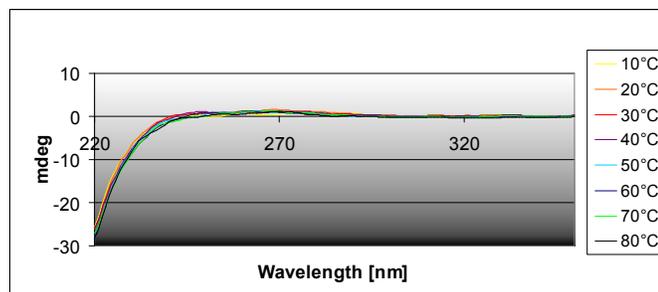


**Figure 5:** Temperature dependent CD spectra of dodecamer duplex sequence ON-4 recorded in cacodylate buffer.

This was though, in contrast to previous study with PNA werein a single amino acid was able to transfer the chirality to the macromolecule. We could also infer that the hybridization of ON-3 was not solely due to the presence of 4 flanking deoxy residues.

It is well known that some DNA polymerases are capable of template directed primer extension using acyclic triphosphate analogues on deoxynucleotides templates.[61,62] That a deoxy ribose backbone primer can be extended using non-natural nucleoside triphosphate and a non-natural oligonucleotide can be extended using natural nucleoside triphosphate, even in the absence of stable duplex formation between the two, is of great significance with regard to the evolution of nucleic acids and its progenitors.[63-66] Our next attempt thus was to study if the achiral acyclic adenine containing DEANA- triphosphate (Figure 6A) is recognized by DNA polymerases in a template assisted, enzymatic, primer extension assay (Figure SI-1). This study was undertaken less in view of the potential prebiotic occurrence of DEANA, but because of its potential adaptability to contemporary biology. We however were not able to detect the template directed primer extension by our acyclic DEANA triphosphate. The results and discussion are provided in the supporting information (Figures-SI-1 and SI-2).



Template (**T**):3'- GTCCTT TGTCGATACTGTCCCC-5'

Primer (**P**):   5'- CAGGAAACAGCTATGAC-3'
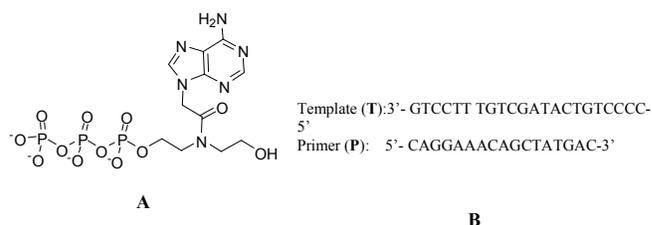
**Figure 6:** (**A**) DEANA triphosphate used in the study; (**B**) Template (**T**) and Primer (**P**) sequence used in the Therminator[TM] DNA polymerase catalysed primer extension assay

## Conclusions

We have constructed two alternative base pairing systems i.e. DEANA and TEANA which are acyclic and achiral, lack sugar units but, have phosphodiester backbone linking the monomeric units. The polymers contain eight bonds between repeating units instead of six bonds found in natural DNA/RNA backbone. The oligonucleotides formed by the DEANA were found to be labile under the standard base deprotection conditions. However, we did manage to obtain deoxynucleotides end capped oligo in low yields for some initial structural studies. The CD study of the dodecamer and modeling of the acyclic nucleotide polymers indicate that a right handed B-type helix could (theoretically) be formed by the mechanism of chiral induction. Further, the model of DEANA oligonucleotide showed that the helical parameters of a DEANA duplex are similar to the PNA parameters. In the model, the phosphate groups are stretched outside, away from the hydrophobic core. We also subjected DEANA triphosphate to a template assisted, polymerase catalyzed incorporation assay. At least, with Therminator[TM] DNA polymerase, we were not able to detect the incorporation of DEANA nucleotide in DNA. Further assays with other mutant polymerases might provide for the potential use of our analogues in future. The synthesis of monomeric units of both DEANA and TEANA are plausible under prebiotic conditions.

## Experimental

### General experimental procedures

All moisture-sensitive reactions were performed under nitrogen atmosphere by using oven-dried glassware. Dry solvents were purchased from commercial sources and used as such. Reactions were monitored by TLC (pre-coated silica gel plate F254, Merck). Column chromatography: Merck Kieselgel 60 (70–230 mesh); flash chromatography: Merck Kieselgel 60 (230–400 mesh). $^{1}$H and $^{13}$C NMR spectra were recorded on NMR spectrometers operating at 300 or 500 MHz and 75 or 125 MHz, respectively. DEANA triphosphate (Figure 6A) was synthesized according to the one-pot synthesis described by Ludwig for natural triphosphates[67] and purified on a DEAE Sephadex-A25 column using a triethylammonium bicarbonate gradient and lyophilized for further use.

### Molecular Modeling (Creation of a static model of a DEANA double stranded helix).

The peptide nucleotide structure (1PUP, 1.7Ang,)[68] was used as a template to construct a 3D model of a DEANA duplex. First, a DEANA dinucleotide ApT was drawn using Chemdraw 10 and 3D coordinates were generated and optimized by Chem3D.[69] This structure was then superimposed onto a PNA dinucleotide unit (with A and T bases), using a flexible superposition program where the torsion angles are optimized while superimposing the conformations onto the rigid PNA dinucleotide. The scoring was done in a similar way as in the SEAL program for steric and electronic alignment.[70] Despite the fact that the DEANA backbone has one more atom than the PNA backbone, a very good superposition of the bases was obtained. This is possible because the phosphate groups are pushed outside. Using that new conformation a DEANA hexamer duplex model with sequence CGTACG.CGTACG was constructed by combining, editing the model pdb file, eventually changing some bases by superposition to get the correct sequence using Quatfit.[71] The model was submitted to a vacuum energy minimization in the ff99bsc0 force field using the Amber 12 software.[72] The force field had to be patched with some gaff force field parameters because of the non-conventional DEANA residues.[73] Charges on atoms of the phosphate groups were adjusted to preserve correct total charges of the nucleotides.

### Synthesis of the DEANA monomers with A, G, C and T bases.

The ethyl esters of Thymine-1-yl acetic acid (**1a**), (2.2g, 10mmol); Cytosine-1-yl acetic acid (**1b**), (1.9g, 10mmol), Adenine-9-yl acetic acid (**1c**), (2.2 g, 10mmol) and Guanine-9-yl acetic acid (**1d**) (2.3g, 10mmol) were refluxed with 2,2'-iminodiethanol (10 eq) overnight in 50 ml of dry ethanol. After overnight reflux the reaction was cooled and filtered. The white precipitate was washed with cold ethanol and diethyl ether to obtain N,N-Bis-(2-hydroxyethyl)-2-(thymin-1-yl)acetamide (**2a**); N,N-Bis-(2-hydroxyethyl)-2-(cytosine-1-yl)acetamide (**2b**); N,N-Bis-(2-hydroxyethyl)-2-(adenine-9-yl)acetamide (**2c**); and N,N-Bis-(2-hydroxyethyl)-2-(guanine-1-yl)acetamide

(**2d**); respectively as white solids, analogous to the synthesis of N,N-Bis-(2-hydroxyethyl)-2-(thymin-1-yl)acetamide[39] reported in the literature.

**N,N-bis(2-hydroxyethyl)-2-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)acetamide (2a):** Yield=95%. $^{1}$H NMR (500MHz; DMSO-$d_6$) δ: 7.27 (1H, d, J= 1.1 Hz), 4.59 (2H, s) 3.56 (2H, appt, J = 6 Hz), 3.46 (2H, *app*t, 6Hz), 3.40 (2H, *app*t, J =6.0 Hz) 3.33 (2H, *app*t, J=6Hz) 1.73 (3H, d, J= 1.1Hz) ; $^{13}$C (125MHz; DMSO-$d_6$) δ: 167.4, 164.9, 151.4, 142.8, 108.5, 59.1, 58.8, 50.0, 48.9, 48.5, 12.2; HRMS (ESI+): calcd for $C_{11}H_{18}N_3O_5$ [M + H]$^{+}$ 272.1240, found 272.1241

**N,N-bis(2-hydroxyethyl)-2-(4-amino-2-oxopyrimidin-1(2H)-yl)-acetamide (2b):** Yield=92%. $^{1}$H NMR (300MHz; DMSO-$d_6$) δ: 7.4 (1H, d, J =7.26Hz), 5.64 (1H, d, J =7.26Hz), 4.58 (2H, *b*s)3.65-3.27 (8H, m); $^{13}$C--(75MHz; DMSO-$d_6$) δ : 167.7, 166.2, 156.0, 147.1, 93.0, 59.1, 58.7, 49.9, 49.1, 48.8; HRMS (ESI+): calcd for $C_{10}H_{16}N_4O_4Na$ [M + Na]$^{+}$ 279.1063, found 279.1062

**2-(6-amino-9H-purin-9-yl)-N,N-bis(2-hydroxyethyl)acetamide (2c):** Yield=96%. $^{1}$H NMR (300MHz;DMSO-$d_6$) δ: 8.23 (1H, s), 8.11 (1H, s), 5.31 (2H, s), 3.64 (2H, *app*t, J =6Hz) 3.54 (2H, *app*t, J =6Hz), 3.47 (2H, *app*t, J =6Hz), 3.36 (2H, *app*t, J =6 Hz); $^{13}$C (75MHz; DMSO-$d_6$) δ : 167.1, 155.9, 152.7, 150.2, 142.6, 118.3, 59.1, 58.7, 50.2, 48.9, 44.4; HRMS (ESI+): calcd for $C_{11}H_{17}N_6O_3$ [M + H]$^{+}$ 281.13565 found 281.1355

**2-(2-amino-6-oxo-1H-purin-9(6H)-yl)-N,N-bis(2-hydroxyethyl)acetamide (2d):** Yield=85%. $^{1}$H NMR (300MHz;DMSO-$d_6$) δ: 7.52 (1H, s) 4.95 (2H,s) 3.6-3.3 (8H,m); $^{13}$C (75MHz; DMSO-$d_6$) δ : 166.6, 156.8, 153.4, 151.6, 138.3, 115.9, 58.8, 58.5, 49.8, 48.5, 43.8; HRMS (ESI+): calcd for $C_{11}H_{17}N_6O_3$ [M + H]$^{+}$ 281.13565 found 281.1355

**N-(2-(bis(4-methoxyphenyl)(phenyl)methoxy)ethyl)-N-(2-hydroxyethyl)-2-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)acetamide (4a):** In a dry round bottomed flask was suspended, 2b (2.7 g, 10 mmol) in dry pyridine (5 ml), and co-evaporated twice. The residue was suspended again in dry pyridine (50ml) and cooled in ice bath. To this pre-cooled solution was added 4,4′-Dimethoxytriphenylmethyl chloride (3.38 g, 10 mmol) and the reaction stirred overnight at room temperature. After completion the reaction was quenched with saturated aqueous NaHCO$_3$ and extracted thrice with DCM (50 mL). The organic layer were pooled, dried, evaporated and chromatographed over silica (0-5% MeOH/DCM, v/v: 5% pyridine) to obtain the product 4a as white foam (3.78g, 6.6mmol; yield=66%).

$^{1}$H NMR: Mix. of rotamers; (300MHz; CDCl$_3$/D$_2$O) δ: 7.39-7.15 (20H, *m*) 6.93 (d, H-6, *min*, J =1.27Hz), 6.87-6.80 (6H, *m*), 6.43 (*d*, H-6, *maj*, J =1.3Hz), 4.73 (*s*, CH$_2$, *maj*), 4.62 (s,CH$_2$, *min*) 3.79 (s, 2*OCH3, *maj*), 3.7 (s, 2*OMe, *min*), 3.67 (*app*t, CH$_2$, J = 5.83Hz, 4.73Hz), 3.60 (*app*t, CH$_2$, J = 6.74Hz, 4.73Hz), 3.40 (*app*t, CH$_2$, J = 5.83Hz, 4.73Hz), 3.32 (*app*t, CH$_2$, J = 5.83Hz, 4.73Hz), 1.88 (d,T-CH$_3$, *min*, J =1.2Hz), 1.76 (d, T-CH$_3$, *maj*, J =1.2Hz); $^{13}$C (75MHz; DMSO-$d_6$) δ: 167.0, 166.7, 164.4, 164.3, 158.1, 158.0, 151.0, 144.9, 144.6, 142.2, 141.9, 135.6, 135.5, 129.7, 128.8, 128.2, 127.9, 127.8, 127.6,

126.8, 126.6, 125.3, 113.2, 107.9, 86.2, 85.5, 61.2, 58.8, 58.5, 55.0, 49.8, 48.1, 47.9, 46.0, 21.0, 11.9; HRMS (ESI+): calcd for $C_{32}H_{35}N_3O_7Na$ 596.2367 found 596.2375

**2-(N-(2-(bis(4-methoxyphenyl)(phenyl)methoxy)ethyl)-2-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)acetamido)ethyl 2-cyanoethyl diisopropylphosphoramidite (5a):** 4a (5.0g, 8.7mmol) was dissolved in dry DCM (40ml). DIPEA (2.3ml, 13mmol) was added and the reaction cooled to 0° C. To this was added 2-cynoethyl N, N diisopropylphosphoramidochloridite (2.3 mL, 10.4 mmol), and the mixture was stirred for 15-30 min. at room temperature. After completion, the reaction was quenched by addition of MeOH (1ml) and solvent evaporated. The residue was then treated with 10% aq.$NaHCO_3$ solution and extracted thrice with DCM (100ml). The organic layers were pooled, dried, evaporated and chromatographed over silica (0-50% Acetone/Hexane, v/v: 2% $Et_3N$) to obtain the product 5a (4.7g, 6.0mmol; yield= 70%) as white foam.

$^{31}$P NMR (121MHz; $CDCl_3$) δ: 148.5, 147.9; HRMS (ESI+): calcd for $C_{41}H_{53}N_5O_8P$ [M + H]$^+$ 774.3639, found 774.3626

**N-(1-(2-(bis(2-hydroxyethyl)amino)-2-oxoethyl)-2-oxo-1,2-dihydropyrimidin-4-yl)benzamide (3b):** In a round bottomed flask, 2b (2.56g, 10mmol) was suspended in dry pyridine (5ml) and co-evaporated twice. The residue was suspended again in dry pyridine (50ml) and cooled to 0° C in an ice bath. To this pre-cooled suspension was added trimethylsilyl chloride (5.0ml, 40mmol) and the reaction stirred at 0° C for 30 min and then at room temperature for 1 hr. The reaction was cooled again to 0° C and freshly distilled benzoyl chloride (3.0 ml, 25mmol) was added drop-wise to this turbid solution. The reaction was stirred at 0° C for 1 hr and then left to stir at room temperature for 3 hr. After completion, the reaction was cooled and quenched by addition of 25 ml of 50:50 methanol: water mixture and stirred for 1 hr. After the evaporation of the solvent the gummy residue was chromatographed over silica (0-10% MeOH/DCM, v/v: 5% pyridine) to obtain the product 3b (2.8g, 7.7 mmol, yield= 77%) as white powder.

$^{1}$H NMR (300MHz; DMSO-$d_6$) δ: 8.0 (3H, m), 7.62 (1H, d, $J$ =7.0 Hz), 7.51 (2H, m), 7.29 (1H, d, $J$ =7.0Hz), 4.83 (2H, s), 3.62 (2H, m), 3.50 (4H, m), 3.38 (2H, m); $^{13}$C (75MHz; DMSO-$d_6$) δ : 166.8, 163.3, 155.3, 151.4, 133.4, 132.8, 128.5, 95.7, 59.1, 58.7, 50.1, 48.9; HRMS (ESI+): calcd for $C_{17}H_{21}N_4O_5$ [M + H]$^+$ 361.1506 found 361.1505

**N-(1-(2-((2-(bis(4-methoxyphenyl)(phenyl)methoxy)ethyl)(2-hydroxyethyl)amino)-2-oxoethyl)-2-oxo-1,2-dihydropyrimidin-4-yl)benzamide (4b):** 3b (2.5g, 7.0 mmol) was dissolved in dry DMF (40 ml). $Et_3N$ (3.0ml, 21mmol) was added and the reaction cooled to 0° C in an ice bath. To this pre-cooled solution was added 4,4′-Dimethoxytriphenylmethyl chloride (2.3 g, 7.0mmol) and the reaction stirred overnight at room temperature. The reaction was quenched with methanol (1 ml) and the solvent evaporated. The residue was chromatographed over silica (0-10% MeOH/DCM, v/v; 5% pyridine) to obtain the product 4b (3.0 g, 4.55 mmol; yield= 65%) as white foam.

$^{1}$H NMR Mix. of rotamers; (500MHz, DMSO-$d_6$) δ: 8.02 (1H, d, $J$ =7.6Hz), 7.94 (d, $J$ = 6.4Hz), 7.62 (*app*t, $J$ =8.3Hz, 7.2Hz), 7.51 (*app*t, $J$ =8.3 Hz, 7.5Hz), 7.41-7.21 (m), 6.90 (d, $J$ =8.6Hz), 4.88 (s,$CH_2$, *maj*), 4.74 (s, $CH_2$, *min*), 3.73(s, 2*$OCH_3$, *maj*), 3.66-3.54 (m), 3.49 (*app*t, $J$ =6.71Hz, 5.12Hz), 3.43 (m), 3.36 (s, 2*$OCH_3$, *min*), 3.34-3.27 (m), 3.09 (*app*t, $J$ =7.22Hz, 5.46Hz); $^{13}$C (125MHz, DMSO-$d_6$) δ: 167.3, 167.0, 166.6, 163.4, 158.2, 158.1, 155.4, 151.4, 145.0, 144.8, 135.7, 135.6, 133.3, 132.7, 129.8, 129.7, 128.5, 128.0, 127.87, 127.7, 126.7, 113.3, 95.7, 86.3, 85.6, 61.4, 61.3, 59.0, 58.5, 55.1, 50.1, 49.9, 48.5, 47.5, 46.3; HRMS (ESI+): calcd for $C_{38}H_{39}N_4O_7$ [M + H]$^+$ 663.2813 found 663.2813

**2-(2-(4-benzamido-2-oxopyrimidin-1(2H)-yl)-N-(2-(bis(4-methoxyphenyl)(phenyl) methoxy)ethyl)acetamido)ethyl 2-cyanoethyl diisopropyl-phosphoramidite (5b):** 4b (3.0g, 4.55mmol) was dissolved in dry THF (25ml). DIPEA (4.0ml, 22.75mmol) was added and the reaction cooled to 0° C. To this was added 2-cynoethyl N,N diisopropyl-phosphoramido-chloridite (1.5 mL, 6.8 mmol), and the mixture was stirred for 15-30 min. at room temperature. After completion, the reaction was quenched by addition MeOH (1 ml) and the solvent evaporated. The residue was chromatographed over silica (0-50% Acetone/Hexane, v/v: 2% $Et_3N$) to obtain the product 5b (2.4g, 2.8mmol; yield= 62%) as white foam.

$^{31}$P- NMR (121MHz; $CDCl_3$) δ: 148.7, 147.8; HRMS (ESI+): calcd for $C_{47}H_{55}N_6O_8P$ Na [M + Na]$^+$ 885.3711, found 885.3719

**N-benzoyl-N-(9-(2-(bis(2-hydroxyethyl)amino)-2-oxoethyl)-9H-purin-6-yl)benzamide (3c):** In a round bottomed flask was suspended 2c (2.8g, 10mmol) in dry pyridine and co-evaporated twice. The residue was suspended again in dry pyridine (50ml) and cooled to 0° C in an ice bath. To this pre-cooled suspension was added Trimethylsilyl chloride (5.0ml, 40mmol) and the reaction stirred at 0° C for 30 min and then at room temperature for 1 hr. The reaction was cooled again to 0° C and freshly distilled benzoyl chloride (3.0 ml, 25mmol) was added drop-wise to this turbid solution. The reaction was stirred at 0° C for 1 hr and then left to stir at room temperature for 3 hr. After completion, the reaction was cooled and quenched by addition of 10 ml of 50:50 methanol:water mixture and left to stir for 1 hr. The solvent was evaporated and the gummy residue was chromatographed over silica (0-10% MeOH/DCM, v/v: 5% pyridine) to obtain the product 3c (4.0g, 8.2mmol; yield=82%) as white powder.

$^{1}$H NMR (300MHz; $CDCl_3$) δ: 8.55 (1H, s), 8.24 (1H,s), 8.05 (2H, m), 7.82 (4H, m),7.44 (3H, m), 7.31 (5H, m), 5.20 (2H, s), 3.62 (4H, bs), 3.41 (4H, bs); $^{13}$C (75MHz; $CDCl_3$) δ: 172.6, 167.4, 153.6, 152.0, 151.3, 147.2, 134.0, 133.2, 129.4, 128.8, 59.9, 59.7, 51.1, 50.1, 44.9.

HRMS (ESI+): calcd for $C_{25}H_{25}N_6O_5$ [M + H]$^+$ 489.1880 found 489.1881

**N-benzoyl-N-(9-(2-((2-(bis(4-methoxyphenyl)(phenyl) methoxy)ethyl)(2-hydroxyethyl)amino)-2-oxoethyl)-9H-purin-6-yl)benzamide (4c):** 3c (4.0g, 8.2mmol) was dissolved in dry DMF (40ml). $Et_3N$ (3.5ml, 25mmol) was added and the reaction cooled to 0° C in an ice bath. To this pre-cooled

solution was added 4,4′-Dimethoxytriphenylmethyl chloride (3.3 g, 9.8 mmol) and the reaction stirred overnight at room temperature. After completion, the reaction was quenched with methanol (1ml) and the solvent evaporated. The residue was chromatographed over silica (5% pyridine/0-10% MeOH/DCM, v/v) to obtain the product 4c (4.0g, 5.0 mmol; yield=62%) as yellowish foam.

$^1$H NMR (300MHz; CDCl$_3$) δ: 8.56 (s, *min*), 8.54, s, *maj*), 8.14 (s, *min*), 7.95 (s, *maj*), 7.47-7.17 (several multiplets), 6.87-6.77 (several multiplets), 5.26 (s, CH$_2$, *maj*), 5.21 (s, CH$_2$, *min*), 3.76 (s, OMe, *maj*), 3.75 (s, OMe, *min*), 3.70-3.59 (several multiplets), 3.55-3.38 (several multiplets), 3.33 (*app*t, *J* =5.83Hz, 4.77Hz); $^{13}$C (75MHz; CDCl$_3$) δ: 172.4, 167.2, 166.3, 158.8, 158.6, 153.6, 153.5, 152.0, 151.5, 146.6, 146.2, 144.7, 144.4, 135.9, 135.4, 134.2, 134.1, 133.0, 130.2, 130.0, 129.5, 128.8, 128.2, 128.1, 127.9, 127.3, 126.9, 126.8, 113.4, 113.3, 87.5, 86.7, 61.8, 61.0, 59.9, 55.4, 55.3, 53.5, 50.7, 49.6, 49.0, 46.8, 44.6; HRMS (ESI+): calcd for C$_{46}$H$_{43}$N$_6$O$_7$ [M + H]$^+$ 791.3187, found 791.3201

**2-(2-(6-(N-benzoylbenzamido)-9H-purin-9-yl)-N-(2-(bis(4-methoxyphenyl)(phenyl)methoxy)ethyl)acetamido)ethyl 2-cyanoethyl diisopropyl phosphoramidite (5c):** In a dry round bottomed flask, 4c (3.0g, 3.8mmol) was dissolved in dry THF (20ml). DIPEA (3.3ml, 19mmol) was added and the reaction cooled to 0° C. 2-cynoethyl N,N-diisopropylphosphoramidochloridite (1.0 mL, 4.56 mmol) was added and the reaction stirred for 15-30 min. at room temperature. After completion, the reaction was quenched by addition of 1ml of MeOH and solvent evaporated. The residue was chromatographed over silica (0-60% Acetone/Hexane, v/v; 2% Et$_3$N) to obtain the product 5c (2.82g, 2.85mmol; yield=75%) as yellowish foam.

$^{31}$P- NMR (121MHz; CDCl$_3$) δ: 149.1, 148.0; HRMS (ESI+): calcd for C$_{55}$H$_{60}$N$_8$O$_8$P [M + H]$^+$ 991.4266, found 991.4266

**N-(9-(2-(bis(2-hydroxyethyl)amino)-2-oxoethyl)-6-oxo-6,9-dihydro-1H-purin-2-yl)isobutyramide (3d):** In a round bottomed flask, 2d (3.0g, 11mmol) was suspended in dry pyridine (1 ml) and co-evaporated twice. The residue was suspended again in dry pyridine (50ml) and cooled to 0° C in an ice bath. To this pre-cooled suspension was added trimethylsilyl chloride (5.0ml, 40mmol) and the reaction stirred at 0° C for 30 min and then at room temperature for 1 hr. The reaction was cooled again to 0° C and freshly distilled isobutyryl chloride (2.6 ml, 25mmol) was added drop-wise to this turbid solution. The reaction was stirred at 0° C for 1 hr and then left to stir at room temperature for 3 hr. After completion, the reaction was cooled and quenched by addition of 10 ml of 50:50 methanol: water mixture. The solvent was evaporated and the residue was chromatographed over silica (0-10% MeOH/DCM, v/v: 5% pyridine) to obtain the product 3d (3.6g, 9.9mmol; yield=90%) as white powder.

$^1$H NMR (300MHz; DMSO-$d_6$) δ: 7.83(s, 1H), 5.11 (2H, s), 3.64 (2H, bs), 3.53-3.47 (4H, m), 3.38-3.34 (2H, m) 2.77 (1H,m), 1.10 (6H, d, *J* =6.7Hz); $^{13}$C (75MHz; DMSO-$d_6$) δ: 180.1, 166.3, 154.9, 149.3, 147.8, 140.7, 119.5, 58.8, 58.5,

49.8, 48.5, 44.5, 34.6, 18.8; HRMS (ESI+): calcd for C$_{15}$H$_{23}$N$_6$O$_5$ [M + H]$^+$ 367.1724 found 367.1723

**2-(N-(2-(bis(4-methoxyphenyl)(phenyl)methoxy)ethyl)-2-(2-isobutyramido-6-oxo-1H-purin-9(6H)-yl)acetamido)ethyl 2-cyanoethyl diisopropylphosphoramidite (5d):** In a round bottomed flask was suspended 3d (3.0g, 8.0mmol) in dry DMF (40 ml). Et$_3$N (3.5ml, 24mmol) was added and the reaction cooled to 0° C in an ice bath. To this pre-cooled solution was added 4,4′-Dimethoxytriphenylmethyl chloride (3.25 g, 9.6 mmol) and the reaction stirred overnight at room temperature. After the completion, the reaction was quenched with methanol and the solvent evaporated. The residue was taken up for the next reaction. This crude 4d was dissolved in dry THF. DIPEA (4.0ml, 22.5 mmol) was added and the reaction cooled to 0° C. To this was added 2-cynoethyl N,N-diisopropylphosphoramidochloridite (1.20 mL, 5.4 mmol) drop-wise, and the reaction stirred for 15-30 min. at room temperature. After completion, the reaction was quenched by addition of MeOH (1 ml) and the solvent evaporated. The residue was chromatographed over silica (0-60% Acetone/Hexane-2% Et$_3$N, v/v) to give the product 5d (2.5g, 3.0mmol, yield=66%) as yellowish foam.

$^{31}$P- NMR (121MHz; CDCl$_3$) δ: 148.7, 147.0; HRMS (ESI+): calcd for C$_{45}$H$_{58}$N$_8$O$_8$P [M + H]$^+$ 869.4109, found 869.4114

**Synthesis of the TEANA monomer with adenine nucleobase.**

**9-(2-bromoethyl)-9H-purin-6-amine (6a):** Adenine (10g, 74mmol) and K$_2$CO$_3$ (20g, 148mmol) were suspended in dry DMF (500ml) and to this was added dibromoethane (12ml, 148mmol) drop-wise and left to stir at room temperature overnight. After completion, the reaction was filtered and the solvent evaporated *in vacuo*. The solid residue left after evaporation of the solvent was washed with water and ether and dried to obtain the product 6a (13g, 55mmol, yield=74%) as yellow orange solid.

$^1$H NMR (300MHz; DMSO-$d_6$) δ: 8.16 (1H, s), 8.14 (1H, s), 4.56 (2H, *app*t, *J* =7.37 Hz, 6.03Hz), 3.93 (2H, *app*t, *J* =7.37 Hz, 6.03Hz); $^{13}$C (75MHz; DMSO-$d_6$) δ: 156.0, 152.5, 149.5, 118.7, 44.7, 31.6; HRMS (ESI+): calcd for C$_7$H$_9$BrN$_5$ [M + H]$^+$ 242.0036, found 242.0036

**9-(2-(bis(2-(tert-butyldimethylsilyloxy)ethyl)amino)ethyl)-9H-purin-6-amine (6b):** The compound 6a (10g, 41mmol) was suspended in 100ml dry dioxane. To this was added bis-silyl protected ethanolamine (13g, 41mmol). Next CsCO$_3$ (13.35g, 41mmol was added and the reaction refluxed overnight. After completion, the reaction was filtered and the filtrate evaporated. The residue was chromatographed on a silica column (0-60% Acetone/Hexane-2% Et$_3$N, v/v) to obtain the product 6b (12g, 26mmol, yield= 64%)

$^1$H NMR (500 MHz; DMSO-$d_6$) δ: 8.10 (1H,s), 8.0 (1H, s), 4.12 (2H, *app*t, *J* =6.5 Hz, 5.9Hz), 3.41 (4H, , *app*t, *J* =6.9 Hz, 6.4Hz), 2.89 (2H, , *app*t, *J* =6.5 Hz, 5.9Hz), 2.61 (4H, , *app*t, *J* =7.1 Hz, 6.3Hz), 0.8 (18H, s), 0.07 (12H); $^{13}$C (125 MHz; DMSO-$d_6$) δ: 155.9, 152.2, 149.5, 141.4, 118.7, 61.5, 56.5, 54.3, 41.7, 25.8, 17.9, -5.4; HRMS (ESI+): calcd for C$_{23}$H$_{47}$N$_6$O$_2$Si$_2$ [M + H]$^+$ 495.3293, found 495.3289

**N-benzoyl-N-(9-(2-(bis(2-(tert-butyldimethylsilyloxy)ethyl)amino)ethyl)-9H-purin-6-yl)benzamide (6c):** Compound **6b** (10g, 20mmol) was co-evaporated thrice with 5ml dry pyridine and dissolved in the same (50ml) and cooled in an ice bath to 0 °C. Freshly distilled benzoyl chloride (5.0ml, 45mmol) was added drop-wise to this cooled solution and the reaction left to stir for 6 hr. After the disappearance of the starting material the reaction was quenched with NaHCO$_3$ in ice. The aq. layer was extracted with DCM. The organic layers were pooled, dried, evaporated and the residue chromatographed on silica (0-50% EtOAc/Hexane v/v) to obtain the product **6c** (11g, 16mmol, yield=82%)

$^1$H NMR (300MHz; CDCl$_3$) δ: 8.65 (1H, s), 8.36 (1H, s), 7.85 (4H, m), 7.45-7.33 (6H, m), 4.29 (2H, *appt*, *J* =6.7 Hz, 5.5Hz), 3.57 (4H, *appt*, *J* =6.7 Hz, 5.9Hz), 3.03 (2H, *appt*, *J* =6.7 Hz, 5.6Hz), 2.70 (4H, *appt*, *J* =6.8 Hz, 6.0 Hz),0.86 (18H, s), 0.02 (12H, s); $^{13}$C (75MHz; CDCl$_3$) δ : 172.4, 153.4, 151.9, 151.5, 146.4, 134.4, 132.8, 129.5, 128.7, 127.5, 61.9, 57.3, 55.2, 42.9, 26.0, 18.3, -5.2; HRMS (ESI+): calcd for C$_{37}$H$_{55}$N$_6$O$_4$Si$_2$ [M + H]$^+$ 703.3817, found 703.3808

**N-benzoyl-N-(9-(2-(bis(2-hydroxyethyl)amino)ethyl)-9H-purin-6-yl)benzamide (6d):**

In a plastic falcon tube **6c** (7g, 10mmol) was dissolved in 10ml dry THF and cooled. To this was added triethylamine trihydrofluoride (3.2ml, 20mmol) and the reaction stirred at room temperature overnight. The reaction was quenched by addition of solid silica and chromatographed over silica (10-60% Acetone/Hexane-2% Et$_3$N, v/v) to obtain the product **6d** (4.0g, 8.5mmol, yield=85%)

$^1$H NMR (300MHz; DMSO-$d_6$) δ: 8.65 (1H, s), 8.62 (1H, s), 7.78 (5H, m), 7.47 (5H, m), 4.33 (2H, *appt*, *J* =6.8 Hz, 5.6Hz), 3.22 (4H, *appt*, *J* =6.8 Hz, 6.2Hz), 2.91 (2H, *appt*, *J* =6.5 Hz, 5.9Hz), 2.54 (4H, m); $^{13}$C (75MHz; CDCl$_3$) δ : 172.7, 153.6, 151.8, 151.0, 149.8, 148.6, 137.2, 134.0, 133.9, 129.5, 129.4, 128.8, 127.4, 124.7, 59.4, 56.7, 54.2, 42.8; HRMS (ESI+): calcd for C$_{25}$H$_{27}$N$_6$O$_4$ [M + H]$^+$ 475.2088, found 475.2086

**N-benzoyl-N-(9-(2-((2-(bis(4-methoxyphenyl)(phenyl)methoxy)ethyl)(2-hydroxyethyl)amino)ethyl)-9H-purin-6-yl)benzamide (6e):**

In a round bottomed flask, **6d** (4.0g, 8.5mmol) was co-evaporated twice with dry pyridine and dissolved 50ml of the same. 4,4′-Dimethoxytriphenylmethyl chloride (1.7 g, 5.1 mmol) was added and the reaction left to stir for 2 hr at room temperature and quenched by addition of methanol (Reaction was only 40-50% complete at this time). The solvent was evaporated and the residue treated with ice cold NaHCO$_3$. The aq. layer was extracted with DCM. The organic layers were pooled, dried, evaporated and chromatographed over silica to obtain the product **6e** (3.2g, 4.2mmol, yield=50%)

$^1$H NMR (300MHz; CDCl$_3$) δ: 8.58 (1H, s), 8.05 (1H, s), 7.83 (4H, m), 7.47-7.18 (15H, m), 4.24 (2H, *appt*, *J* =6.6 Hz, 5.9Hz), 3.34 (2H, *appt*, *J* =6.6 Hz, 5.0Hz), 3.17 (2H, *appt*, *J* =6.0 Hz, 5.6Hz), 2.98 (2H, *appt*, *J* =7.0 Hz, 6.0Hz), 2.69 (2H, *appt*, *J* =6.3 Hz, 5.6Hz), 2.62 (2H, *appt*, *J* =7.0 Hz, 5.0Hz); $^{13}$C (75MHz; CDCl$_3$) δ :171.9, 157.9, 153.0, 151.3, 150.4, 147.5,

145.0, 135.8, 133.5, 133.2, 129.5, 129.1, 128.8, 128.6, 128.4, 127.7, 127.6, 126.6, 126.5, 113.1, 85.6, 61.7, 59.2, 56.3, 54.9, 45.7, 42.1; HRMS (ESI+): calcd for C$_{46}$H$_{44}$N$_6$O$_6$Na [M + Na]$^+$ 799.3214, found 799.3214

**2-((2-(6-(N-benzoylbenzamido)-9H-purin-9-yl)ethyl)(2-(bis(4-methoxyphenyl)(phenyl)methoxy)ethyl)amino)ethyl 2-cyanoethyl diisopropyl phosphoramidite (6f):** In a round bottomed flask, **6d** (1.0g, 1.28 mmol) was dissolved in dry THF and cooled to 0° C in ice bath. 2-cynoethyl N,N-diisopropylphosphoramidochloridite (0.3 mL, 1.28mmol) was then added drop-wise to this pre-cooled solution and left to stir in the ice bath for 15 min and then ice bath was removed. After completion, the reaction was quenched by addition of methanol and the solvent evaporated. The residue was chromatographed on silica to obtain the product **6f** (0.65g, 0.66mmolyield=52%) as white foam. $^{31}$P- NMR (121MHz; CDCl$_3$) δ: 147.7; HRMS (ESI+): calcd for C$_{55}$H$_{62}$N$_8$O$_7$P [M + H]$^+$ 977.4473, found 977.4471

## Notes and references

$^a$ Medicinal Chemistry, KU Leuven, Minderbroedersstraat 10, B-3000 Leuven, Belgium.
$^b$ Organic Materials Chemistry and Bioorganic Chemistry, Institute or Chemistry, University of Osnabrück, Barharastrasse 7, 49069 Osnabrück, Germany.

Electronic Supplementary Information (ESI) available: General experimental procedures, details of primer extension assay, Tm profiles of oligos, the hairpin model. See DOI: 10.1039/b000000x/

(1)  Orgel, L. E. *Trends Biochem. Sci.* **1998**, *23*, 491-495.
(2)  Miller, S. L.; Urey, H. C. *Science* **1959**, *130*, 245-251.
(3)  Miller, S. L. *J. Am. Chem. Soc.* **1955**, *77*, 2351-2361.
(4)  Joyce, G. F. *Nature (London)* **1989**, *338*, 217-24.
(5)  House, C. H. *Astrobiology* **2003**, *3*, 245-247.
(6)  McCollom, T. M. *Annu. Rev. Earth Planet. Sci.* **2013**, *41*, 207-229.
(7)  Engelhart, A. E.; Hud, N. V. *Cold Spring Harb Perspect Biol* **2010**, *2*, a002196.
(8)  Orgel, L. E. *Crit. Rev. Biochem. Mol. Biol.* **2004**, *39*, 99.
(9)  Sowerby, S. J.; Petersen, G. B. *Astrobiology* **2002**, *2*, 231-239.
(10)  Chen, I. A.; Nowak, M. A. *Acc. Chem. Res.* **2012**, *45*, 2088-2096.
(11)  Nielsen, P. E. *Orig. Life Evol. Biosph.* **1993**, *23*, 323-327.
(12)  Eschenmoser, A. *Science (Washington, D. C.)* **1999**, *284*, 2118-2124.

(13) Eschenmoser, A. *Orig. Life Evol. Biosphere* **2004**, *34*, 277-306.

(14) Eschenmoser, A. *Chimia* **2005**, *59*, 836-850.

(15) Eschenmoser, A. *Tetrahedron* **2007**, *63*, 12821-12844.

(16) Nelson, K. E.; Levy, M.; Miller, S. L. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 3868-3871.

(17) Larralde, R.; Robertson, M. P.; Miller, S. L. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92*, 8158-60.

(18) Keefe, A.; Miller, S. *J. Mol. Evol.* **1995**, *41*, 693-702.

(19) Cleaves, J. H. *Astrobiology* **2002**, *2*, 403-415.

(20) Tohidi, M.; Orgel, L. E. *J. Mol. Evol.* **1989**, *28*, 367-373.

(21) Nelsestuen, G. *J. Mol. Evol.* **1980**, *15*, 59-72.

(22) Cheikh, A.; Orgel, L. *J. Mol. Evol.* **1990**, *30*, 315-321.

(23) De Clercq, E.; Sakuma, T.; Baba, M.; Pauwels, R.; Balzarini, J.; Rosenberg, I.; Holy, A. *Antiviral Res.* **1987**, *8*, 261-72.

(24) De Clercq, E.; Holy, A. *Nat. Rev. Drug Discov.* **2005**, *4*, 928-940.

(25) Schneider, K. C.; Benner, S. A. *J. Am. Chem. Soc.* **1990**, *112*, 453-455.

(26) Schwartz, A. W.; Orgel, L. E. *Science* **1985**, *228*, 585-587.

(27) Schwartz, A. W.; Visscher, J.; Bakker, C. G.; Niessen, J. *Orig. Life Evol. Biosphere* **1987**, *17*, 351-357.

(28) Joyce, G. F.; Schwartz, A. W.; Miller, S. L.; Orgel, L. E. *Proc. Natl. Acad. Sci. U. S. A.* **1987**, *84*, 4398-402.

(29) Zhang, S.; Switzer, C.; Chaput, J. C. *Chem. Biodiversity* **2010**, *7*, 245-258.

(30) Vandendriessche, F.; Augustyns, K.; Aerschot, V.; Busson, R.; Hoogmartens, J.; Herdewijn, P. *Tetrahedron* **1993**, *49*, 7223-7238.

(31) Van der Laan, A. C.; Stromberg, R.; van Boom, J. H.; Kuyl-Yeheskiely, E.; Efimov, V. A.; Chakhmakhcheva, O. G. *Tetrahedron Lett.* **1996**, *37*, 7857-7860.

(32) Li, P.; Sun, J.; Su, M.; Yang, X.; Tang, X. *Org. Biomol. Chem.*, *12*, 2263-2272.

(33) Kehler, J.; Henriksen, U.; Vejbjerg, H.; Dahl, O. *Bioorg. Med. Chem.* **1998**, *6*, 315-322.

(34) Peyman, A.; Uhlmann, E.; Wagner, K.; Augustin, S.; Breipohl, G.; Will, D. W.; Schaefer, A.; Wallmeier, H. *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 2636-2638.

(35) Sutherland, J. D.; Whitfield, J. N. *Tetrahedron* **1997**, *53*, 11493-11527.

(36) Miller, S. L. *Nat. Struct. Biol.* **1997**, *4*, 167-169.

(37) Dickens, J. E. I., W. M.; Ohishi, M.; Ikeda, M.; Ishikawa,S.; Nummelin, A.; Hjalmarson, Ã *Astrophys. J.* **1997**, *489*, 753-757.

(38) Dueholm, K. L.; Egholm, M.; Behrens, C.; Christensen, L.; Hansen, H. F.; Vulpius, T.; Petersen, K. H.; Berg, R. H.; Nielsen, P. E.; Buchardt, O. *J. Org. Chem.* **1994**, *59*, 5767-5773.

(39) Hadj-Bouazza, A.; Teste, K.; Colombeau, L.; Chaleix, V.; Zerrouki, R.; Kraemer, M.; Sainte Catherine, O. *Nucleos. Nucleot. Nucleic Acids* **2008**, *27*, 439-448.

(40) Fkyerat, A.; Demeunynck, M.; Constant, J. F.; Michon, P.; Lhomme, J. *J. Am. Chem. Soc.* **1993**, *115*, 9952-9959.

(41) Drew, H. R.; Wing, R. M.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R. E. *Proc. Natl. Acad. Sci. U. S. A.* **1981**, *78*, 2179-2183.

(42) Patel, D. J.; Kozlowski, S. A.; Marky, L. A.; Broka, C.; Rice, J. A.; Itakura, K.; Breslauer, K. J. *Biochemistry* **1982**, *21*, 428-436.

(43) Tereshko, V.; Minasov, G.; Egli, M. *J. Am. Chem. Soc.* **1999**, *121*, 3590-3595.

(44) Tereshko, V.; Minasov, G.; Egli, M. *J. Am. Chem. Soc.* **1999**, *121*, 470-471.

(45) Fratini, A. V.; Kopka, M. L.; Drew, H. R.; Dickerson, R. E. *J. Biol. Chem.* **1982**, *257*, 14686-14707.

(46) Dickerson, R. E.; Drew, H. R. *J. Mol. Biol.* **1981**, *149*, 761-786.

(47) Drew, H. R.; Dickerson, R. E. *J. Mol. Biol.* **1981**, *151*, 535-556.

(48) Kool, E. T. *Chem. Rev. (Washington, D. C.)* **1997**, *97*, 1473-1487.

(49) Herdewijn, P. *Liebigs Ann.* **1996**, 1337-1348.

(50) Murayama, K.; Tanaka, Y.; Toda, T.; Kashida, H.; Asanuma, H. *Chem. Eur. J.*, *19*, 14151-14158.

(51) Pallan, P. S.; Lubini, P.; Bolli, M.; Egli, M. *Nucleic Acids Res.* **2007**, *35*, 6611-6624.

(52) He, W.; Hatcher, E.; Balaeff, A.; Beratan, D. N.; Gil, R. R.; Madrid, M.; Achim, C. *J. Am. Chem. Soc.* **2008**, *130*, 13264-13273.

(53) Ratilainen, T.; Holmén, A.; Tuite, E.; Haaima, G.; Christensen, L.; Nielsen, P. E.; Nordén, B. *Biochemistry* **1998**, *37*, 12331-42.

(54) Yeh, J. I.; Pohl, E.; Truan, D.; He, W.; Sheldrick, G. M.; Du, S.; Achim, C. *Chem. - Eur. J.* **2010**, *16*, 11867-11875.

(55) Froeyen, M.; Morvan, F.; Vasseur, J. J.; Nielsen, P.; Van Aerschot, A.; Rosemeyer, H.; Herdewijn, P. *Chemistry & Biodiversity* **2007**, *4*, 803-17.

(56) Rasmussen, H.; Liljefors, T.; Petersson, B.; Nielsen, P. E.; Kastrup, J. S. *J. Biomol. Struct. Dyn.* **2004**, *21*, 495-502.

(57) Kozlov, I. A.; Orgel, L. E.; Nielsen, P. E. *Angew. Chem., Int. Ed.* **2000**, *39*, 4292-4295.

(58) Sforza, S.; Haaima, G.; Marchelli, R.; Nielsen, P. E. *Eur. J. Org. Chem.* **1999**, 197-204.

(59) Wittung, P.; Eriksson, M.; Lyng, R.; Nielsen, P. E.; Norden, B. *J. Am. Chem. Soc.* **1995**, *117*, 10167-73.

(60) Wittung, P.; Nielsen, P. E.; Buchardt, O.; Egholm, M.; Norden, B. *Nature (London)* **1994**, *368*, 561-3.

(61) Heuberger, B. D.; Switzer, C. *J. Am. Chem. Soc.* **2008**, *130*, 412-413.

(62) Chen, J. J.; Tsai, C.-H.; Cai, X.; Horhota, A. T.; McLaughlin, L. W.; Szostak, J. W. *PLoS One* **2009**, *4*, e4949.

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

(64) Horhota, A. T.; Szostak, J. W.; McLaughlin, L. W. *Org. Lett.* **2006**, *8*, 5345-5347.

(65) Chaput, J. C.; Szostak, J. W. *J. Am. Chem. Soc.* **2003**, *125*, 9274-9275.

(66) Chaput, J. C.; Ichida, J. K.; Szostak, J. W. *J. Am. Chem. Soc.* **2003**, *125*, 856-857.

(67) Ludwig, J. *Acta Biochim. Biophys. Acad. Sci. Hung.* **1981**, *16*, 131-133.

(68) Rasmussen, H.; Kastrup, J. S.; Nielsen, J. N.; Nielsen, J. M.; Nielsen, P. E. *Nat. Struct. Biol.* **1997**, *4*, 98-101.

(69) ChemDraw; 10.0 ed. CambridgeSoft.

(70) Kearsley, S. K.; Smith, G. M. *Tetrahedron Comput. Methodol.* **1990**, *3*, 615-633.

(71) Heisterberg, D. J. In *The Quatfit program, the CCL archive* 1990.

(72) Salomon-Ferrer, R.; Case, D. A.; Walker, R. C. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2013**, *3*, 198-210.

(73) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *Journal of Computational Chemistry* **2004**, *25*, 1157-74.