



HAL
open science

Enzymatic synthesis of biphenyl-DNA oligonucleotides

Pascal Röthlisberger, Fabienne Levi-Acobas, Christian Leumann, Marcel Hollenstein

► **To cite this version:**

Pascal Röthlisberger, Fabienne Levi-Acobas, Christian Leumann, Marcel Hollenstein. Enzymatic synthesis of biphenyl-DNA oligonucleotides. *Bioorganic and Medicinal Chemistry*, 2020, 28 (11), pp.115487. 10.1016/j.bmc.2020.115487 . pasteur-02625707

HAL Id: pasteur-02625707

<https://pasteur.hal.science/pasteur-02625707>

Submitted on 6 Jul 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - ShareAlike 4.0 International License

Graphical Abstract

Enzymatic synthesis of biphenyl-DNA oligonucleotides

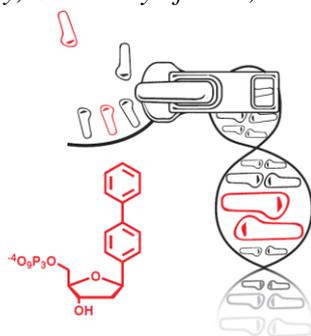
Leave this area blank for abstract info.

Pascal R othlisberger,^{a,b} Fabienne Levi-Acobas,^{a,b} Christian J. Leumann,^c and Marcel Hollenstein^{a,b*}

^a Institut Pasteur, Department of Structural Biology and Chemistry, Laboratory for Bioorganic Chemistry of Nucleic Acids, CNRS UMR 3523, 28, rue du Docteur Roux, 75724 Paris Cedex 15, France.

^b Institut Pasteur, Department of Genome and Genetics, Paris, France

^c Department of Chemistry and Biochemistry, University of Bern, Freiestrasse 3, CH-3012 Bern, Switzerland



Enzymatic synthesis of biphenyl-DNA oligonucleotides

Pascal Röthlisberger,^{a,b} Fabienne Levi-Acobas,^{a,b} Christian J. Leumann,^c and Marcel Hollenstein^{a,b*}

^aInstitut Pasteur, Department of Structural Biology and Chemistry, Laboratory for Bioorganic Chemistry of Nucleic Acids, CNRS UMR 3523, 28, rue du Docteur Roux, 75724 Paris Cedex 15, France.

^bInstitut Pasteur, Department of Genome and Genetics, Paris, France

^cDepartment of Chemistry and Biochemistry, University of Bern, Freiestrasse 3, CH-3012 Bern, Switzerland

* Corresponding author. Tel.: +33-1-44-38-94-66; e-mail: marcel.hollenstein@pasteur.fr

ARTICLE INFO

Article history:

Received

Received in revised form

Accepted

Available online

Keywords:

DNA polymerases

modified nucleic acids

nucleoside triphosphates

biaryls

hydrophobic bases

ABSTRACT

The incorporation of nucleotides equipped with C-glycosidic aromatic nucleobases into DNA and RNA is an alluring strategy for a number of practical applications including fluorescent labelling of oligonucleotides, expansion of the genetic alphabet for the generation of aptamers and semi-synthetic organisms, or the modulation of excess electron transfer within DNA. However, the generation of C-nucleoside containing oligonucleotides relies mainly on solid-phase synthesis which is quite labor intensive and restricted to short sequences. Here, we explore the possibility of constructing biphenyl-modified DNA sequences using enzymatic synthesis. The presence of multiple biphenyl-units or biphenyl residues modified with electron donors and acceptors permits the incorporation of a single **dBphMP** nucleotide. Moreover, templates with multiple abasic sites enable the incorporation of up to two **dBphMP** nucleotides, while TdT-mediated tailing reactions produce single-stranded DNA oligonucleotides with four biphenyl residues appended at the 3'-end.

2009 Elsevier Ltd. All rights reserved.

1. Introduction

DNA is a fascinating biopolymer that mainly serves as the repository of the genetic information. The well-known double helical structure adopted by DNA arises through specific Watson-Crick interactions and packing forces between the four canonical nucleobases A, C, G, and T. However, a combination of organic chemistry and polymerase engineering have not only permitted to broaden the horizon of applications of DNA to fields that markedly deviate from its natural role[1-12] but also to expand the genetic alphabet by one[13-16] or multiple[17] orthogonal, artificial base pairs. Unnatural base pairs (UBPs) that have been used in this context mainly consist of large aromatic residues that interact *via* hydrophobic and packing forces,[18] nucleobases with alternative hydrogen bond patterns,[17] and nucleoside analogs capable of interacting through shape complementarity.[19, 20] More recently, artificial metal base pairs have been proposed as alternative scaffolds for the construction of UBPs.[21-28] Such an expansion of the genetic alphabet has important practical repercussions since semi-synthetic organisms capable of replicating DNA containing artificial nucleotides[29] and even transcribing modified mRNA into proteins[14] could be created and aptamers with enhanced binding properties were identified.[13, 20, 30, 31] In this context,

the zipper-like interstrand stacking motif (Fig. 1A) is an alternative recognition motif based on the interaction of C-nucleotides equipped with bipyridine (Bipy) or biphenyl (**dBph**) moieties within a DNA duplex (Fig. 1B).[32-34] The inclusion of four biphenyl base pairs in opposing positions in short DNA duplexes gave rise to important thermal stabilities (ΔT_m of up to 4.4°C per base pair).[34] Interestingly, up to 7 consecutive biphenyl base pairs can be incorporated into DNA without causing any decrease in duplex stability[32] and an NMR investigation concluded that stacking interactions were the main driving force behind the impressive stability of these UBPs.[35] The zipper-like motif thus appears to be an alluring candidate for applications such as electron transport through nucleic acids[36-38] or an expansion of the genetic code.[39] Besides biphenyl-modified DNA, other C-nucleosides equipped with aromatic residues play central roles in numerous applications including the induction of important conformational transitions,[40] relaying electron transport through duplexes,[37, 38, 41-43] assisting the homogenous assembly of aromatic carbon layers,[44] or labeling of oligonucleotides with fluorophores.[45-49] However, solid-phase synthesis of these analogs and biphenyl-DNA particularly is quite labor intensive and often complicated by the reduced solubility and the tendency to aggregate. Therefore, alternative methods such as enzymatic synthesis of **dBph** containing

oligonucleotides would be highly desirable.[32] Here, we explored the possibility of using enzymatic synthesis to generate oligonucleotides containing multiple **dBph** nucleotides. To this end, we have investigated whether 1. the formation of a zipper-like interstrand stacking motif could favor the successive incorporation of **dBph** nucleotides during primer extension (PEX) reactions and 2. the presence of multiple abasic sites on a template strand could be used to trigger the incorporation of multiple C-nucleotides into DNA. In this context, we demonstrate that the presence of three **dBph** residues on the template permits the incorporation of one modified nucleotide into DNA and that the introduction of abasic sites can be used to generate oligonucleotides containing up to two modified **dBph** analogs. We also show that the appendage of electron donors and acceptors on the distal ring of the biphenyl moiety can be used to improve the efficiency of enzymatic synthesis of DNA modified with aromatic C-nucleotides. Taken together, these results show that the enzymatic synthesis of oligonucleotides containing biphenyl C-nucleotide analogs is possible and could be extended to other zipper-like motifs[50-54] and potentially for the construction of randomized libraries containing UBPs for SELEX experiments.[30, 55-57]

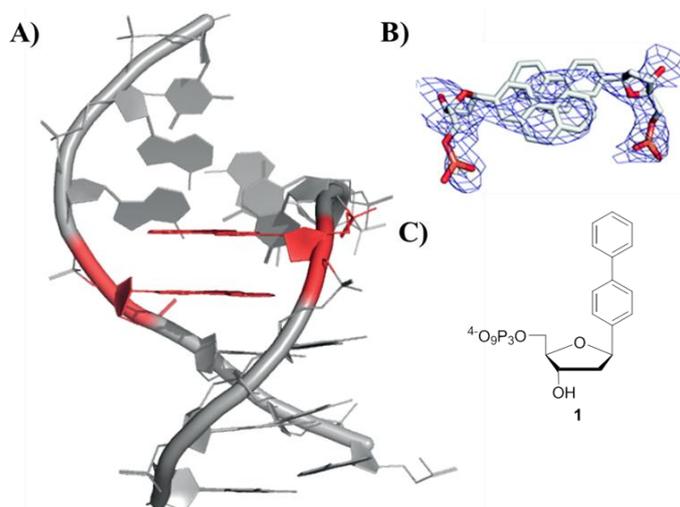
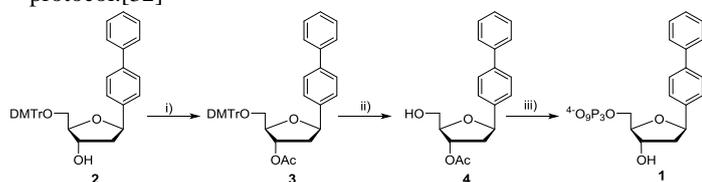


Figure 1. A) and B) details of an X-ray structure of an interstrand-stacking zipper motif (red) adapted from reference [58] and C) chemical structure of **dBphTP 1**.

2. Results and discussion

2.1. Synthesis of the triphosphate and modified oligonucleotides

The synthesis of the modified nucleotide **1** proceeded by application of standard protocols (Scheme 1).[3, 9] Briefly, nucleoside **2**[32, 34] underwent a sequence of protecting group rearrangements which involved acetylation followed by detritylation. The resulting suitably protected nucleoside analog **4** was then triphosphorylated by application of the Ludwig-Eckstein conditions[59] which after anion exchange HPLC (IE-HPLC) afforded **dBphTP 1** in good yields (48%). The corresponding phosphoramidite building block for solid-phase DNA synthesis was obtained by application of a known protocol.[32]



Scheme 1. Synthesis of nucleoside triphosphate **dBphTP 1**. Reagents and conditions: i) Ac₂O, Et₃N, DMAP, pyridine, rt, 2.5h, 95%; ii) DCA, DCM, rt, 30 min, 81%; iii) a) 2-chloro-1,3,2-benzodioxaphosphorin-4-one, pyridine, dioxane, rt, 45 min; (b) (nBu₃NH)₂H₂P₂O₇, DMF, nBu₃N, rt, 45 min; (c) I₂, pyridine, H₂O, rt, 30 min; (d) NH₃(aq.), rt, 1.5 h, 48% (4 steps).

Natural and modified oligonucleotides were either purchased or prepared according to standard protocols for automated DNA synthesis on a 1 μmol scale in the trityl-off mode, respectively (see Table 1 and ESI).[32, 34] The modified nucleotides were introduced *via* the corresponding phosphoramidites (**5-11**, Figure 2) obtained as described previously.[60-62]

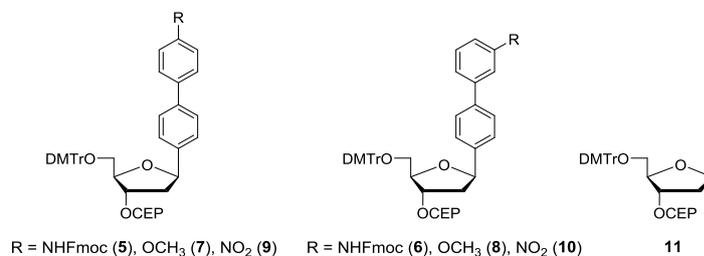


Figure 2. Chemical structures of the modified biphenyl-C-nucleoside (**5-10**) and abasic site nucleoside phosphoramidite **11** used for the synthesis of modified oligonucleotides.

Table 1. Primers and templates used for primer extension reactions.

Name	Sequence (shown in the 5' to 3' direction) ^a
P1	5'-TAC GAC TCA CTA TAG CCT C
T1	5'-GGA GBphG AGG CTA TAG TGA GTC GTA
T2	5'-GGBph BphBphG AGG CTA TAG TGA GTC GTA
T3	5'-GGA GXG AGG CTA TAG TGA GTC GTA
T4	5'-GGA G5G AGG CTA TAG TGA GTC GTA
T5	5'-GGA G6G AGG CTA TAG TGA GTC GTA
T6	5'-GGA G7G AGG CTA TAG TGA GTC GTA
T7	5'-GGA G8G AGG CTA TAG TGA GTC GTA
T8	5'-GGA G9G AGG CTA TAG TGA GTC GTA
T9	5'-GGA G10G AGG CTA TAG TGA GTC GTA
T10	5'-GGA GφG AGG CTA TAG TGA GTC GTA
T11	5'-GGφφφG AGG CTA TAG TGA GTC GTA
T12	5'-GGA GCG AGG CTA TAG TGA GTC GTA

^a Numbers ranging from **5** to **10** represent the modified nucleotides shown in Figure 2 and **φ** = **11**.

2.2. Evaluation of enzymatic synthesis via a zipper-like interstrand stacking motif

With **dBphTP 1** at hand, we proceeded to evaluate the possibility of incorporating this C-nucleotide into DNA by enzymatic synthesis and eventually construct a zipper-like interstrand stacking motif. Therefore, we performed PEX reactions using different polymerases (Phusion, Hemo KlenTaq, Taq, *Bst*, Q5 DNA polymerase, Terminator, Vent (*exo*), *Sulfolobus* DNA polymerase IV (Dpo4), Deep Vent, Phi29, and

the Klenow fragment of DNA polymerase I *exo*⁻ (Kf *exo*⁻) along with **dBphTP 1** as only substrate. We first evaluated the incorporation efficiency using a 24-nucleotide long template **T1** (Table 1) containing a single **dBph** unit immediately downstream of the end of the 18-mer primer **P1**. [27] This initial screen revealed that the modified nucleoside triphosphate was not readily accepted as a substrate since no elongation of the primer was observed (Fig. S1A). A similar observation was made with each of the natural dNTPs (data not shown) as well as with the abasic site nucleotide **dφTP 11** even though this analog was shown to be readily incorporated opposite templating dA units (Fig. S1B). [62] We hypothesized that this poor incorporation efficiency mirrors the trend observed in thermal stability of duplexes where a single **dBph-dBph** base pair caused a significant decrease in T_m . [32] Hence, we assessed whether the presence of multiple **dBph** units on the template would spur the enzymatic formation of the zipper-like motif. We thus carried out PEX reactions with template **T3** that contains three consecutive modified nucleotides immediately downstream of the end of primer **P1** (Fig. 3). This analysis reveals that the polymerase Dpo4 incorporated a single **dBphMP** opposite a modified nucleotide in moderate yields (~50%) while other polymerases such as Hemo KlenTaq and *Bst* only poorly recognized the modified nucleotide as a substrate since only ~10% of the elongated primer was formed which is sometimes accompanied by exonucleolytic degradation of the primer (see e.g. lanes 1 and 6). A similar trend but with different polymerases was observed with **dφTP 11** since the n+1 product formed in ~50% yield with *Bst* and ~10% with Hemo KlenTaq (Fig. 3).

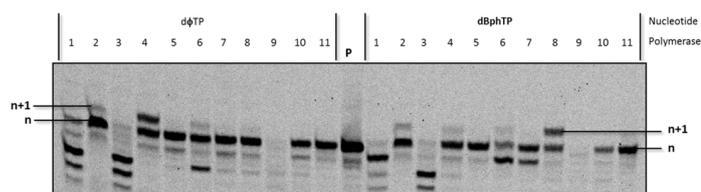


Figure 3. Gel analysis (PAGE 20%) of the products from primer extension reactions with primer **P1**, template **T2**, and the modified nucleotides **dφTP** and **dBphTP**. All the reaction mixtures contained 200 μM of nucleotides and 2 U of the DNA polymerases. The reaction mixtures were incubated at the corresponding temperatures in 1X reaction buffer for 3 h. **P** indicates the unreacted primer. Polymerases used: 1. Phusion; 2. Hemo KlenTaq; 3. Q5 DNA polymerase; 4. *Bst*; 5. Taq; 6. Therminator; 7. Vent (*exo*⁻); 8. Dpo4; 9. Deep Vent; 10. Phi29; 11. Kf *exo*⁻.

2.3. Effect of electron-donating and -withdrawing groups on enzymatic DNA synthesis

After a first evaluation of the possibility of an enzymatic construction of a zipper-like motif using the *C*-nucleoside analog **dBph**, we next questioned whether a modulation of the π electron density on the biphenyl moieties could have an impact on the incorporation efficiency of the modified nucleotide. Indeed, the inclusion of electron-withdrawing groups on the aromatic system was shown to significantly increase the T_m value of the resulting duplexes while the presence of donating groups led to thermoneutral duplexes (mixed duplexes containing acceptor and donor moieties also led to increased T_m values). [61] In order to evaluate whether the nature as well as the position of substituents on the biphenyl moiety could favor the incorporation of a single **dBph** moiety into DNA we synthesized templates **T4-9** that contain single biphenyl-nucleosides equipped with OMe, NH₂, and NO₂ units in the distal ring in both *para* and *meta* orientations (Fig. 2 and Table 1). PEX reactions were carried out

with primer **P1**, **dBphTP**, and various DNA polymerases (Fig. 4 and 5). This analysis clearly reveals that various polymerases including Dpo4 and Kf *exo*⁻ readily incorporate the modified nucleotide opposite biphenyl units equipped with a *para*-amino unit (template **T3**; Fig. 4A and Table S1). In addition, PEX reactions conducted with templates containing *meta*-nitro and *para*-methoxy modifications (**T8** (Fig. 5C) and **T5** (Fig. 4B), respectively) led to the formation of the expected n+1 product albeit in lesser yields and with fewer polymerases. Hence, the presence of electron-donors in the *para*-position or electron-withdrawing groups in the *meta*-position of the distal ring appears to be a means to improve the incorporation efficiency of the modified **dBph** nucleotide into DNA. On the other hand, enzymatic incorporation of the modified nucleotide was markedly reduced when the substitution pattern was reversed i.e. with electron-donors in the *meta*-position and electron-withdrawing groups in the *para*-position (see Fig. 4C and Fig. 5A and B and Table S1). Notwithstanding, partial formation of the n+1 extended primer product was formed in the presence of these templates which is not the case with the unmodified **Bph**-containing template **T1**. The PEX reactions carried out with Terminator also seem to follow this trend even though partial exonucleolytic degradation of the primer can be observed. (Fig. 4 and Fig. 5)

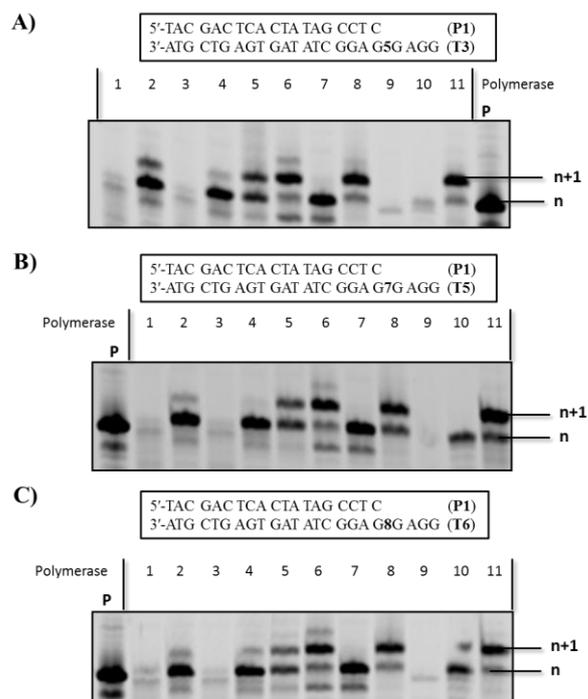


Figure 4. Gel analysis (PAGE 20%) of the products of PEX reactions with **dBphTP**, primer **P1**, and templates A) **T3**, B) **T5**, and C) **T6**. All the reaction mixtures contained 200 μM of nucleotides and 2 U of the DNA polymerases. The reaction mixtures were incubated at the corresponding temperatures in 1X reaction buffer for 3 h. **P** indicates the unreacted primer. Polymerases used: 1. Phusion; 2. Hemo KlenTaq; 3. Q5 DNA polymerase; 4. Taq; 5. *Bst*; 6. Therminator; 7. Vent (*exo*⁻); 8. Dpo4; 9. Deep Vent; 10. Phi29; 11. Kf *exo*⁻.

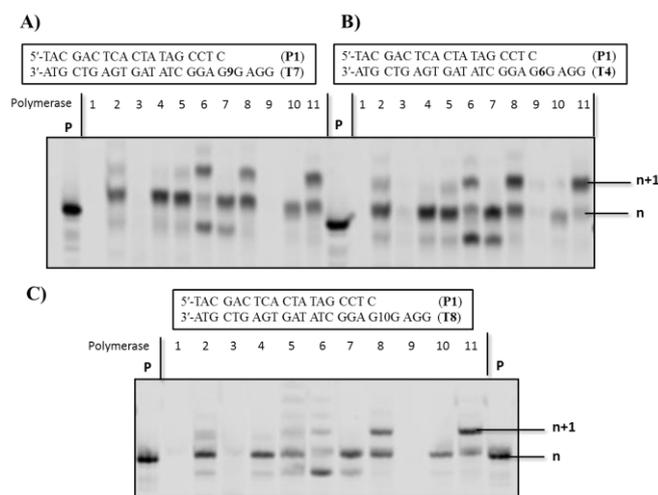


Figure 5. Gel analysis (PAGE 20%) of the products of PEX reactions with **dBphTP**, primer **P1**, and templates A) **T7**, B) **T4**, and C) **T8**. All the reaction mixtures contained 200 μM of nucleotides and 2 U of the DNA polymerases. The reaction mixtures were incubated at the corresponding temperatures in 1X reaction buffer for 3 h. **P** indicates the unreacted primer. Polymerases used: 1. Phusion; 2. Hemo KlenTaq; 3. Q5 DNA polymerase; 4. Taq; 5. *Bst*; 6. Therminator; 7. Vent (*exo*⁻); 8. Dpo4; 9. Deep Vent; 10. Phi29; 11. Kf *exo*⁻.

2.4. Enzymatic construction of oligonucleotides with multiple **dBph** units

Nucleotides equipped with non-hydrogen bonding aromatic moieties have been shown to be incorporated efficiently by DNA polymerases opposite abasic sites.[63-65] Since solid-phase synthesis of **Bph**-containing oligonucleotides is rather labor intensive, we hypothesized that such sequences could be constructed by the sequential enzymatic incorporation of modified nucleotides opposite abasic sites in PEX reactions. To do so, we synthesized three different templates containing one (**T9**), three (**T10**), or five (**T11**) consecutive abasic sites (Table 1). Gel analysis of PEX reactions with **dBphTP** revealed that the n+1 product was formed in good yields with a few DNA polymerases irrespective of the nature of the template (Fig. 6 and ESI). During reactions with template **T9**, a single **dBphMP** was readily incorporated opposite an abasic site by Therminator, Dpo4, and Kf *exo*⁻, while the PEX reactions with all other polymerases either led to little product formation, partial degradation of the primer, or no incorporation event (Fig. S2A and Fig. 6). Importantly, after the installation of such a **dBph- ϕ** pair, enzymatic DNA synthesis could be resumed by Therminator upon addition of natural nucleoside triphosphates (Fig. S3). Therefore, we propose a two-step protocol for the synthesis of any sequence containing a single **dBphMP** unit: 1. PEX reaction with Therminator, Dpo4, or Kf *exo*⁻ and **dBphTP** on a template containing an abasic site 2. natural DNA synthesis after the installation of the **dBph- ϕ** pair with dNTPs and Therminator. Such a protocol has recently been applied for the synthesis of modified DNazymes using metal base pairs[28] and DNA duplexes containing artificial silver base pairs.[66] In the former, a modified nucleotide was installed opposite natural dA and dT nucleotides and enzymatic synthesis was resumed with a different polymerase (Dpo4) while in the latter, a modified purine nucleotide was used to form an artificial silver base pair and natural DNA synthesis was resumed with Therminator. Our

novel strategy mainly differs from these two first examples by the nature of the templating nucleotide (*i.e.* abasic sites) employed to trigger the incorporation of a C-nucleotide analog into DNA.

In the context of templates containing multiple abasic sites, efficient incorporation of a single **dBph** nucleotide was observed but no products corresponding to primer extended by multiple **dBph** units was formed (Fig. S2B). However, given that some of the polymerases displayed a high propensity at incorporating the modified nucleotide (high yields of product formation and no exonucleolytic degradation of the primer), we questioned whether conditions could be met that would enable the incorporation of multiple **dBph** units. We thus performed PEX reactions with increasing concentrations of modified nucleotide and by supplementing the reaction mixtures with 1 mM Mn^{2+} which is known to relax the substrate specificity of DNA polymerases (Fig. 6).[67-70] Variation of both parameters indeed resulted in higher incorporation efficiencies of **dBphMP** and the formation of the n+2 product (in yields up to ~30%), particularly with the polymerase Dpo4 and additional Mn^{2+} . However, enzymatic synthesis of modified DNA appears to stall after the incorporation of two **dBph** unit since no n+3 product formation could be observed regardless of the nature of the template. For comparison, we also evaluated the incorporation efficiency of natural dATP on templates containing multiple consecutive abasic sites. In PEX reactions with template **T10**, Dpo4 efficiently incorporated two dAMP units and the n+3 product could be observed albeit in poor yields (<10%; Fig. 6A). When template **T11** was used instead, Dpo4 produced a distribution of products while Therminator led to the formation of the n+2 and n+3 products. Thus, extension of primers with dAMP appears to be more efficient than with **dBphTP** but is nonetheless impaired by the presence of multiple abasic sites.

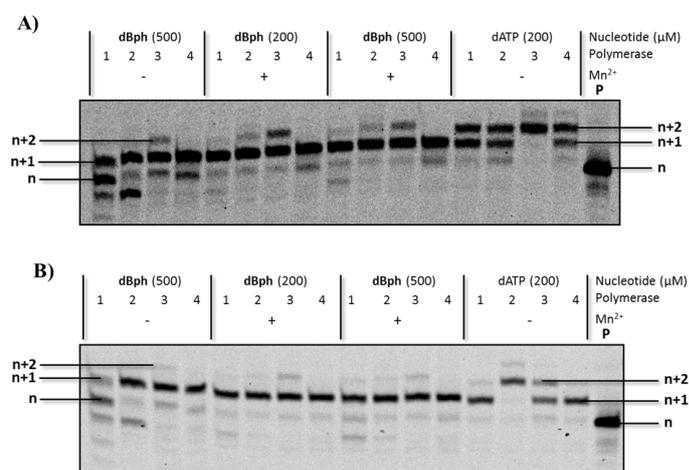


Figure 6. Gel analysis (PAGE 20%) of the products of PEX reactions with **dBphTP**, primer **P1**, and templates A) **T10** and B) **T11** under various experimental conditions. All the reaction mixtures contained 2 U of the DNA polymerases and the reaction mixtures were incubated at the corresponding temperatures in 1X reaction buffer for 3 h. The final concentration of Mn^{2+} was 1 mM. **P** indicates the unreacted primer. Polymerases used: 1. *Bst*; 2. Therminator; 3. Dpo4; 4. Kf *exo*⁻.

Lastly, we evaluated the possibility of generating oligonucleotides equipped with **dBph** tails. To do so, we employed the X family polymerase terminal deoxynucleotidyl transferase (TdT) which incorporates nucleotides at the 3'-termini of ssDNA oligonucleotides in a template-independent manner.[71] Tailing reactions were carried out under different experimental conditions where the nature of the metal cofactors

(Mn²⁺, Mg²⁺, and Co²⁺), the influence of the initiator (*i.e.* primer) concentrations (2 and 20 μM) and the reaction time on the tailing reactions were evaluated (Fig. 7). When both Mg²⁺ and Co²⁺ were used as cofactors with low initiator concentrations, efficient tailing reactions were observed with the incorporation of up to four **dBph** units. On the other hand, when Mn²⁺ was used as a cofactor, the tailing reactions were not as efficient since the polymerase seemed to stall after the addition of three modified nucleotides. The order of preference in terms of metal cofactors thus follows the trend often observed with modified nucleotides, *i.e.* Co²⁺ ≈ Mg²⁺ > Mn²⁺. [71] Expectedly, decreasing the monomer/initiator ratio negatively impacted the TdT-mediated polymerization reactions. [72] The appendage of four *C*-nucleotides on ssDNA primers during TdT-mediated tailing reactions appears to be a general upper limit for such analogs since similar results had been reported with pyrene-modified nucleotides. [73, 74]

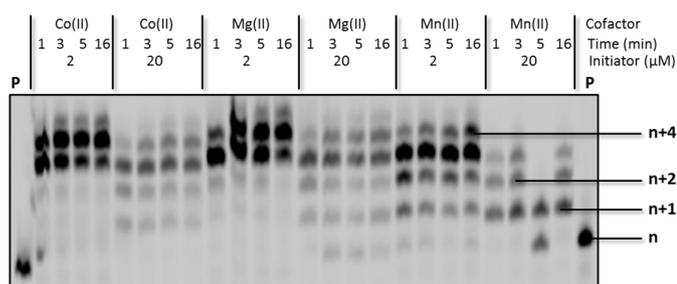


Figure 7. Gel image (PAGE 20%) of the TdT-mediated tailing reactions with **dBphTP** and different initiator concentrations, reaction times, and cofactors. All reactions were performed with primer **P1** and incubated with 10 U of TdT at 37°C. Cofactor concentrations were 1 mM for Mn²⁺ and Mg²⁺ and 0.25 mM for Co²⁺. **P** indicates the unreacted primer.

3. Conclusions

The presence of *C*-nucleosides equipped with aromatic moieties in DNA can be used in a broad range of applications. However, while solid-phase synthesis permits the introduction of nucleoside analogs into DNA and RNA sequences, this method is restricted to rather short sequences and to simple modification patterns. Alternative protocols for the synthesis of oligonucleotides containing aromatic *C*-nucleotides are thus in dire need. [75] Enzymatic synthesis of modified oligonucleotides is an alluring strategy since a single modified nucleotide could be employed for the construction of numerous oligonucleotides potentially containing multiple modified *C*-nucleotides irrespective of the length of the sequence. [76, 77]

Our study shows that the *C*-nucleotide analog **dBphTP** can be used as a building block for the enzymatic construction of modified oligonucleotides under certain experimental conditions. We have shown that the presence of consecutive modified residues on template oligonucleotides can partially improve the incorporation efficiency of **dBphMP** into DNA even though the yields remain modest (~50% conversion was obtained with the Dpo4 polymerase). However, the incorporation efficiency can be further improved by changing the stereoelectronic nature of the distal ring of templating **dBph** nucleotides. Best incorporation efficiency was obtained by introducing a para-amino moiety on the distal ring combined with the Dpo4 or the Kf *exo*⁻ polymerases. Moreover, multiple **dBph** building blocks can be incorporated into DNA by the Terminator, Dpo4, or Kf *exo*⁻ polymerases either by using templates containing abasic sites or by the TdT-tailing reaction which seems to be restricted to only

four incorporation events. We have also devised a two-step enzymatic protocol that permits the synthesis of any sequence containing a **dBph** moiety. This two-step method can potentially be applied to other modified (*e.g.* pyrene-modified nucleotides [63]) or natural (*e.g.* dATP [62]) nucleotides that are known to be incorporated opposite basic sites. Collectively, these results demonstrate the possibility of using polymerases to synthesize **dBph**-containing oligonucleotides and could be expanded to other *C*-nucleotides. This approach can benefit from the use of engineered polymerases [8, 78, 79] and work towards this aim is currently under way in our laboratory.

4. Experimental section

4.1. Materials and methods

All reactions were performed under argon in flame-dried glassware. Anhydrous solvents for reactions were obtained from Sigma Aldrich. Flash chromatography was performed using silica gel (230–400 mesh) from Sigma Aldrich or on a Reveleris Prep system from Büchi. Thin layer chromatography was carried out on pre-coated glass-backed plates of silica gel (0.25 mm, UV₂₅₄) from Macherey-Nagel. All chemicals and solvents used were purchased from Sigma-Aldrich and Alfa Aesar unless stated otherwise. NMR spectra were recorded on a Bruker Avance 400 spectrometer (400.13 MHz for ¹H, 100.62 MHz for ¹³C, and 161.62 MHz for ³¹P) and all spectra were referenced to the signals of the corresponding solvent. Chemical shifts are given in ppm (δ scale) and coupling constants (*J*) in Hz. Assignment of the NMR signals was performed by using a combination of ¹H/¹H-COSY, ¹³C-DEPT-135, and ¹³C/¹H-HMBC experiments. High resolution electrospray ionization (ESI) mass spectra (MS, *m/z*) were recorded on a Waters Q-ToF Micro MS in the positive-ion electrospray ionization (ESI+) mode. Solutions were prepared using 1:1 MeCN/H₂O containing 0.1% formic acid or MeOH/water containing 10 mM ammonium acetate in the case of sensitive compounds. HPLC purification was performed using an Äkta™ pure system (GE Healthcare) equipped with Thermo Scientific™ DNAPac™ PA100 semi preparative ion exchange column (13 μm, 250 x 9.0 mm). Unmodified DNA oligonucleotides were purchased from Microsynth. DNA oligonucleotides containing modified nucleotides were synthesized on an H-8 DNA synthesizer from K&A on a 1 μmol scale. Natural DNA phosphoramidites (dT, dC4bz, dG2DMF, dA6Bz) and solid support (dA6Bz-lcaa-CPG 500Å) were all purchased from ChemGenes. Natural DNA phosphoramidites as well as the modified phosphoramidite were prepared as 0.07 M solutions in MeCN and were coupled using 50 sec and 490 sec steps, respectively. 5-(ethylthio)-1*H*-tetrazole (0.25 M in MeCN) was used as coupling agent. Capping, oxidation, and detritylation were performed using standard conditions. Cleavage from the solid support and deprotection of oligonucleotides was achieved by treatment with concentrated ammonia at 55°C for 16 h. After centrifugation, the supernatants were collected and the resulting solutions were evaporated to dryness on a speed-vac. Crude oligonucleotides were purified by anion exchange HPLC (Dionex - DNAPac PA200). Buffer solutions of 25 mM Tris-HCl in H₂O, pH 8.0 (buffer A) and 25 mM Tris-HCl, 1.25 M NaCl in H₂O, pH 8.0 (buffer B) were used. The purified oligonucleotides were then desalted with SepPack C-18 cartridges. Oligonucleotide concentrations were quantitated by UV spectroscopy using a UV5Nano spectrophotometer (Mettler Toledo). The chemical integrity of oligonucleotides was assessed by UPLC-MS analysis: UPLC was performed on a BEH C18 column (130 Å, 1.7 μm, 2.1 mm x 50 mm) from Waters, installed on an ACQUITY UPLC H-Class System (SQ Detector 2). A Buffer containing 20 mM TEA and 400 mM HFIP in H₂O was used with

a linear gradient from 18 to 31% Methanol within 5 minutes and a flow rate of 0.3 mL/min. All the DNA polymerases (Phusion, Hemo KlenTaq, Taq, *Bst*, Q5, Therminator, Vent (*exo*), Dpo4, Deep Vent, Phi29, and Kf *exo*) were purchased from New England Biolabs as well as the natural dNTPs. Acrylamide/bisacrylamide (29:1, 40%) was obtained from Fisher Scientific. Visualization of PAGE gels was performed by fluorescence imaging using a Storm 860 or a Typhoon Trio phosphorimager with the ImageQuant software (both from GE Healthcare).

4.2. Chemical synthesis of nucleotides

Synthesis of nucleoside analog 3:

The starting material **2** [32, 34] (100mg, 0.17mmol) was dissolved in dry pyridine (2.5 mL) at RT under N₂ and DMAP (24 mg, 0.19 mmol, 1.1eq.) and Et₃N (61 μL, 0.43 mmol, 2.5 eq.) was added. To this solution acetic anhydride (24 μL, 0.26mmol, 1.5 eq.) was added in one portion. The reaction mixture was stirred for 2.5h then concentrated to dryness. The crude oil was extracted in DCM (3 x 10mL) and NaHCO₃ sat (5 mL). The organic layers were combined dried over MgSO₄ and concentrated. The crude product was purified by flash chromatography (Hex/EtOAc 4/1) to yield 102 mg (95%) of the title compound **3**.

R_f (Hex/ EtOAc 4/1) = 0.4. MS C₄₀H₃₈O₆Na⁺ calculated 637.256; observed 637.2585.

¹H NMR (400 MHz, Chloroform-*d*) δ = 7.61 (m, 4H), 7.52 (dd, *J*=7.8, 2.1, 4H), 7.46 (m, 2H), 7.43 – 7.36 (m, 4H), 7.33 – 7.27 (m, 2H), 7.26 – 7.22 (m, 1H), 6.94 – 6.65 (m, 4H), 5.44 – 5.31 (m, 1H), 5.18 (dd, *J*=10.9, 5.0, 1H), 4.23 (d, *J*=2.0, 1H), 3.97 – 3.74 (m, 6H), 3.47 – 3.29 (m, 2H), 2.46 – 2.34 (m, 1H), 2.25 (ddd, *J*=13.6, 11.0, 5.9, 1H), 2.13 (s, 3H).

¹³C NMR (101 MHz, Chloroform-*d*) δ = 170.5, 158.5, 144.9, 140.9, 140.7, 140.1, 136.1, 136.1, 130.2, 128.8, 128.3, 127.8, 127.3, 127.2, 127.1, 126.8, 126.6, 113.1, 86.3, 84.3, 80.2, 64.3, 55.2, 41.6, 21.2.

Synthesis of nucleoside analog 4:

Nucleoside **3** (90mg, 0.14 mmol) was dissolved in dry DCM (9 mL) and then treated with dichloroacetic acid (120μL 1.4mmol, 10eq.). The resulting orange reaction mixture was stirred for 30 min. under N₂ at RT. The reaction mixture was quenched with MeOH and evaporated to dryness. The crude oil was extracted in DCM (3 x 10mL) and NaHCO₃ sat (5 mL). The organic layers were combined dried over MgSO₄ and concentrated. The crude product was purified by flash chromatography (Hex/EtOAc 4/1) to yield 37 mg (81%) of the title compound **4**.

R_f (Hex/ EtOAc 4/1) = 0.15.

MS C₁₉H₂₀O₄Na⁺ calculated 312.1362; observed 637.2585.

¹H NMR (400 MHz, Methanol-*d*₄) δ = 7.65 – 7.56 (m, 4H), 7.54 – 7.48 (m, 2H), 7.47 – 7.40 (m, 2H), 7.35 – 7.30 (m, 1H), 5.28 (ddd, *J*=6.0, 1.8, 0.9, 1H), 5.12 (dd, *J*=11.0, 5.0, 1H), 4.08 (td, *J*=4.4, 2.2, 1H), 3.79 (dd, *J*=4.5, 1.2, 2H), 3.33 (p, *J*=1.7, 1H), 2.35 (ddd, *J*=13.7, 5.0, 1.1, 1H), 2.12 (s, 3H), 2.15 – 2.06 (m, 1H).

¹³C NMR (101 MHz, Methanol-*d*₄) δ = 171.0, 140.8, 140.7, 140.0, 128.4, 127.0, 126.6, 126.5, 126.4, 85.7, 80.2, 76.8, 62.3, 41.1, 19.6.

Synthesis of nucleoside triphosphate 1:

The suitably protected nucleoside **4** (37 mg, 0.11 mmol) was dissolved in dry pyridine (0.2 mL) and dry dioxane (0.4 mL) at RT under N₂. To this clear solution, 2-chloro-1,3,2-benzodioxaphosphorin-4-one (40 mg, 0.18 mmol, 1.5 eq.) were added and the reaction mixture was stirred for 45 min. A solution of tributylammonium pyrophosphate (95 mg, 0.17 mmol, 1.4 eq.) in dry DMF (0.17 mL) and tributylamine (60 μL) was added dropwise and the reaction mixture stirred for another 45 min. The reaction mixture was then oxidized by the addition of iodine (56 mg, 0.21 mmol, 1.8 eq.) in pyridine (0.98 mL) and H₂O (0.02 mL). After 30 min of stirring the excess of iodine was quenched with a sodium thiosulfate (10% w/v in water) and the clear solution was concentrated under reduced pressure at 30°C. The concentrated mixture was treated with ammonium hydroxide 30% (12 mL) for 2h. The yellow suspension was again concentrated under reduced pressure at 30°C. The yellow residue was dissolved in H₂O (2 mL) and precipitated by the addition of NaClO₄ 2% in acetone (12 mL). The crude product was purified by RP-HPLC (30% B in 20 min; Buffer A) TEAB 50mM pH=8; Buffer B) TEAB 50mM pH=8, 50% ACN) to yield 29 mg (48 %) of the pure triphosphate **1**.

MS (MALDI): *m/z* calculated for C₁₇H₁₉O₁₂P₃²⁻: 508.01; observed: 508.24.

¹H NMR (400 MHz, Deuterium Oxide) δ = 7.72 – 7.57 (m, 4H), 7.50 (d, *J*=8.0, 2H), 7.43 (t, *J*=7.7, 2H), 7.38 – 7.31 (m, 1H), 5.16 (dd, *J*=10.5, 5.7, 1H), 4.52 (d, *J*=5.4, 1H), 4.14 (s, 1H), 4.06 (t, *J*=4.8, 2H), 2.26 – 2.08 (m, 2H).

³¹P NMR (162 MHz, Deuterium Oxide) δ = -6.70 (d, *J*=13.8), -9.78 (d, *J*=19.1), -21.28 (t, *J*= 14.6).

4.3. Enzymatic reactions

The 5'-FAM-labelled primer **P1** (30 pmol) was annealed to the appropriate template (45 pmol) in water by heating to 95°C and then gradually cooling to room temperature. The appropriate reaction buffer, the modified triphosphate; and DNA polymerases were added to the annealed oligonucleotides for a total reaction volume of 20 μL. Following incubation at the optimal temperature for the enzyme, the reactions were quenched by adding stop solution (20 μL; formamide (70%), ethylenediaminetetraacetic acid (EDTA; 50 mM), bromophenol (0.1%), xylene cyanol (0.1 %)). The reaction mixtures were subjected to gel electrophoresis in denaturing polyacrylamide gel (20%) containing trisborate-EDTA (TBE) 1x buffer (pH 8) and urea (7M). Visualization was performed by phosphorimaging.

For TdT-tailing reactions, solutions containing 20 or 2 pmol primer **P1**, 1 U μL⁻¹ of TdT, **dBphTP** (200 μM), 1x reaction buffer, and H₂O (for a total reaction volume of 10 μL) was incubated at 37 °C for 5 h. The reaction mixtures were then purified by Nucleospin columns and quenched by the addition of an equal volume of loading buffer. The reaction products were then resolved by electrophoresis (PAGE 20 %) and visualized by phosphorimager analysis.

Acknowledgments

The authors gratefully acknowledge financial support from Institut Pasteur. The authors thank Dr. Luke McKenzie for the recoding of the MALDI spectrum of the modified triphosphate and Ms. Marie Flamme for some gel images.

References and notes

- [1] P. R thlisberger, M. Hollenstein, *Aptamer Chemistry*, *Adv. Drug Deliv. Rev.*, 134 (2018) 3-21.
- [2] A.I. Taylor, V.B. Pinheiro, M.J. Smola, A.S. Morgunov, S. Peak-Chew, C. Cozens, K.M. Weeks, P. Herdewijn, P. Holliger, *Catalysts from synthetic genetic polymers*, *Nature*, 518 (2015) 427-430.
- [3] M. Hocek, *Enzymatic Synthesis of Base-Functionalized Nucleic Acids for Sensing, Cross-linking, and Modulation of Protein-DNA Binding and Transcription*, *Accounts Chem. Res.*, 52 (2019) 1730-1737.
- [4] M. Hollenstein, *Nucleic acid enzymes based on functionalized nucleosides*, *Curr. Opin. Chem. Biol.*, 52 (2019) 93-101.
- [5] A. Goyenvalle, G. Griffith, A. Babbs, S. El Andaloussi, K. Ezzat, A. Avril, B. Dugovic, R. Chaussenot, A. Ferry, T. Voit, H. Amthor, C. Buhr, S. Schurch, M.J.A. Wood, K.E. Davies, C. Vaillend, C. Leumann, L. Garcia, *Functional correction in mouse models of muscular dystrophy using exon-skipping tricyclo-DNA oligomers*, *Nat. Med.*, 21 (2015) 270-275.
- [6] M.R. Dunn, R.M. Jimenez, J.C. Chaput, *Analysis of aptamer discovery and technology*, *Nat. Rev. Chem.*, 1 (2017) 0076.
- [7] C. Gasse, M. Zaarour, S. Noppen, M. Abramov, P. Marliere, S. Liekens, B. De Strooper, P. Herdewijn, *Modulation of BACE1 Activity by Chemically Modified Aptamers*, *ChemBioChem*, 19 (2018) 754-763.
- [8] V.B. Pinheiro, A.I. Taylor, C. Cozens, M. Abramov, M. Renders, S. Zhang, J.C. Chaput, J. Wengel, S.Y. Peak-Chew, S.H. McLaughlin, P. Herdewijn, P. Holliger, *Synthetic Genetic Polymers Capable of Heredity and Evolution*, *Science*, 336 (2012) 341-344.
- [9] M. Flamme, L.K. McKenzie, I. Sarac, M. Hollenstein, *Chemical methods for the modification of RNA*, *Methods*, 161 (2019) 64-82.
- [10] E. Ereemeeva, A. Fikatas, L. Margamuljana, M. Abramov, D. Schols, E. Groaz, P. Herdewijn, *Highly stable hexitol based XNA aptamers targeting the vascular endothelial growth factor*, *Nucleic Acids Res.*, 47 (2019) 4927-4939.
- [11] P. Chidchob, D. Offenbartl-Stiegert, D. McCarthy, X. Luo, J.N. Li, S. Howorka, H.F. Sleiman, *Spatial Presentation of Cholesterol Units on a DNA Cube as a Determinant of Membrane Protein-Mimicking Functions*, *J. Am. Chem. Soc.*, 141 (2019) 1100-1108.
- [12] K.E. Bujold, A. Lacroix, H.F. Sleiman, *DNA Nanostructures at the Interface with Biology*, *Chem*, 4 (2018) 495-521.
- [13] M. Kimoto, R. Yamashige, K. Matsunaga, S. Yokoyama, I. Hirao, *Generation of high-affinity DNA aptamers using an expanded genetic alphabet*, *Nat. Biotechnol.*, 31 (2013) 453-+.
- [14] Y. Zhang, J.L. Ptacin, E.C. Fischer, H.R. Aerni, C.E. Caffaro, K.S. Jose, A.W. Feldman, C.R. Turner, F.E. Romesberg, *A semi-synthetic organism that stores and retrieves increased genetic information*, *Nature*, 551 (2017) 644-647.
- [15] D.A. Malyshev, K. Dhami, T. Lavergne, T.J. Chen, N. Dai, J.M. Foster, I.R. Correa, F.E. Romesberg, *A semi-synthetic organism with an expanded genetic alphabet*, *Nature*, 509 (2014) 385-388.
- [16] D.A. Malyshev, F.E. Romesberg, *The Expanded Genetic Alphabet*, *Angew. Chem. Int. Ed.*, 54 (2015) 11930-11944.
- [17] S. Hoshika, N.A. Leal, M.J. Kim, M.S. Kim, N.B. Karalkar, H.J. Kim, A.M. Bates, N.E. Watkins, H.A. SantaLucia, A.J. Meyer, S. DasGupta, J.A. Piccirilli, A.D. Ellington, J. SantaLucia, M.M. Georgiadis, S.A. Benner, *Hachimoji DNA and RNA: A genetic system with eight building blocks*, *Science*, 363 (2019) 884-887.
- [18] A.W. Feldman, V.T. Dien, R.J. Karadeema, E.C. Fischer, Y.B. You, B.A. Anderson, R. Krishnamurthy, J.S. Chen, L.J. Li, F.E. Romesberg, *Optimization of Replication, Transcription, and Translation in a Semi-Synthetic Organism*, *J. Am. Chem. Soc.*, 141 (2019) 10644-10653.
- [19] I. Singh, R. Laos, S. Hoshika, S.A. Benner, M.M. Georgiadis, *Snapshots of an evolved DNA polymerase pre- and post-incorporation of an unnatural nucleotide*, *Nucleic Acids Res.*, 46 (2018) 7977-7988.
- [20] M. Kimoto, Y.W.S. Lim, I. Hirao, *Molecular affinity rulers: systematic evaluation of DNA aptamers for their applicabilities in ELISA*, *Nucleic Acids Res.*, 47 (2019) 8362-8374.
- [21] J. M ller, *Nucleic acid duplexes with metal-mediated base pairs and their structures*, *Coord. Chem. Rev.*, 393 (2019) 37-47.
- [22] T. Kobayashi, Y. Takezawa, A. Sakamoto, M. Shionoya, *Enzymatic synthesis of ligand-bearing DNAs for metal-mediated base pairing utilising a template-independent polymerase*, *Chem. Commun.*, 52 (2016) 3762-3765.
- [23] E.K. Kim, C. Switzer, *Polymerase Recognition of a Watson-Crick-Like Metal-Mediated Base Pair: Purine-2,6-Dicarboxylate center dot Copper(II)center dot Pyridine*, *ChemBioChem*, 14 (2013) 2403-2407.
- [24] C. Kaul, M. Muller, M. Wagner, S. Schneider, T. Carell, *Reversible bond formation enables the replication and amplification of a crosslinking salen complex as an orthogonal base pair*, *Nat. Chem.*, 3 (2011) 794-800.
- [25] F. Levi-Acobas, P. R thlisberger, I. Sarac, P. Marliere, P. Herdewijn, M. Hollenstein, *On the Enzymatic Formation of Metal Base Pairs with Thiolated and pK(a)-Perturbed Nucleotides*, *ChemBioChem*, (2019) 10.1002/cbic.201900399.
- [26] P. R thlisberger, F. Levi-Acobas, I. Sarac, P. Marliere, P. Herdewijn, M. Hollenstein, *Towards the enzymatic formation of artificial metal base pairs with a carboxy-imidazole-modified nucleotide*, *J. Inorg. Biochem.*, 191 (2019) 154-163.
- [27] P. R thlisberger, F. Levi-Acobas, I. Sarac, P. Marliere, P. Herdewijn, M. Hollenstein, *On the enzymatic incorporation of an imidazole nucleotide into DNA*, *Org. Biomol. Chem.*, 15 (2017) 4449-4455.
- [28] Y. Takezawa, T. Nakama, M. Shionoya, *Enzymatic Synthesis of Cu(II)-Responsive Deoxyribozymes through Polymerase Incorporation of Artificial Ligand-Type Nucleotides*, *J. Am. Chem. Soc.*, 141 (2019) 19342-19350.
- [29] Y.K. Zhang, B.M. Lamb, A.W. Feldman, A.X. Zhou, T. Lavergne, L.J. Li, F.E. Romesberg, *A semisynthetic organism engineered for the stable expansion of the genetic alphabet*, *Proc. Natl. Acad. Sci. U.S.A.*, 114 (2017) 1317-1322.
- [30] L.Q. Zhang, Z.Y. Yang, K. Sefah, K.M. Bradley, S. Hoshika, M.J. Kim, H.J. Kim, G.Z. Zhu, E. Jimenez, S. Cansiz, I.T. Teng, C. Champanhac, C. McLendon, C. Liu, W. Zhang, D.L. Gerloff, Z. Huang, W.H. Tan, S.A. Benner, *Evolution of Functional Six-Nucleotide DNA*, *J. Am. Chem. Soc.*, 137 (2015) 6734-6737.
- [31] L.Q. Zhang, Z.Y. Yang, T.L. Trinh, I.T. Teng, S. Wang, K.M. Bradley, S. Hoshika, Q.F. Wu, S. Cansiz, D.J. Rowold, C. McLendon, M.S. Kim, Y. Wu, C. Cui, Y. Liu, W.J. Hou, K. Stewart, S. Wan, C. Liu, S.A. Benner, W.H. Tan, *Aptamers against Cells Overexpressing Glypican 3 from Expanded Genetic Systems Combined with Cell Engineering and Laboratory Evolution*, *Angew. Chem. Int. Ed.*, 55 (2016) 12372-12375.
- [32] C. Brotschi, G. Mathis, C.J. Leumann, *Bipyridyl- and biphenyl-DNA: A recognition motif based on interstrand aromatic stacking*, *Chem. Eur. J.*, 11 (2005) 1911-1923.
- [33] C. Brotschi, A. Haberli, C.J. Leumann, *A stable DNA duplex containing a non-hydrogen-bonding and non-shape-complementary base couple: Interstrand stacking as the stability determining factor*, *Angew. Chem. Int. Ed.*, 40 (2001) 3012-3014.
- [34] C. Brotschi, C.J. Leumann, *DNA with hydrophobic base substitutes: A stable, zipperlike recognition motif based on interstrand-stacking interactions*, *Angew. Chem. Int. Ed.*, 42 (2003) 1655-1658.
- [35] Z. Johar, A. Zahn, C.J. Leumann, B. Jaun, *Solution structure of a DNA duplex containing a biphenyl pair*, *Chem. Eur. J.*, 14 (2008) 1080-1086.
- [36] P. R thlisberger, V. Kaliginediand, C.J. Leumann, *Modulation of Excess Electron Transfer through LUMO Gradients in DNA Containing Phenanthrenyl Base Surrogates*, *Chem. Eur. J.*, 23 (2017) 2022-2025.
- [37] J. Frommer, B. Karg, K. Weisz, S. Muller, *Preparation and characterization of pyrene modified uridine derivatives as potential electron donors in RNA*, *Org. Biomol. Chem.*, 16 (2018) 7663-7673.
- [38] L. Antusch, N. Gass, H.A. Wagenknecht, *Elucidation of the Dexter-Type Energy Transfer in DNA by Thymine-Thymine Dimer*

- Formation Using Photosensitizers as Artificial Nucleosides, *Angew. Chem. Int. Ed.*, 56 (2017) 1385-1389.
- [39] C. Brotschi, C.J. Leumann, RNA duplexes with biphenyl substituents as base replacements are less stable than DNA duplexes, *Chem. Commun.*, (2005) 2023-2025.
- [40] M. Nakamura, T. Takada, K. Yamana, Controlling Pyrene Association in DNA Duplexes by B- to Z-DNA Transitions, *ChemBioChem*, 20 (2019) 2949-2954.
- [41] P. Roethlisberger, F. Wojciechowski, C.J. Leumann, Enhancement of Excess Electron Transfer Efficiency in DNA Containing a Phenothiazine Donor and Multiple Stable Phenanthrenyl Base Pairs, *Chem. Eur. J.*, 19 (2013) 11518-11521.
- [42] D.G. Singleton, R. Hussain, G. Siligardi, P. Kumar, P.J. Hrdlicka, N. Berova, E. Stulz, Increased duplex stabilization in porphyrin-LNA zipper arrays with structure dependent exciton coupling, *Org. Biomol. Chem.*, 14 (2016) 149-157.
- [43] E. Batzner, Y. Liang, C. Schweigert, A.N. Unterreiner, H.A. Wagenknecht, Acceleration of Long-Range Photoinduced Electron Transfer through DNA by Hydroxyquinolines as Artificial Base Pairs, *ChemPhysChem*, 16 (2015) 1607-1612.
- [44] K.M. Chan, W. Xu, H. Kwon, A.M. Kietrys, E.T. Kool, Luminescent Carbon Dot Mimics Assembled on DNA, *J. Am. Chem. Soc.*, 139 (2017) 13147-13155.
- [45] D.L. Wilson, E.T. Kool, Fluorescent Probes of DNA Repair, *ACS Chem. Biol.*, 13 (2018) 1721-1733.
- [46] D.L. Wilson, A.A. Beharry, A. Srivastava, T.R. O'Connor, E.T. Kool, Fluorescence Probes for ALKBH2 Allow the Measurement of DNA Alkylation Repair and Drug Resistance Responses, *Angew. Chem. Int. Ed.*, 57 (2018) 12896-12900.
- [47] M. Nakamura, M. Fukuda, T. Takada, K. Yamana, Highly ordered pyrene pi-stacks on an RNA duplex display static excimer fluorescence, *Org. Biomol. Chem.*, 10 (2012) 9620-9626.
- [48] A.R. Rovira, A. Fin, Y. Tor, Expanding a fluorescent RNA alphabet: synthesis, photophysics and utility of isothiazole-derived purine nucleoside surrogates, *Chem. Sci.*, 8 (2017) 2983-2993.
- [49] W. Xu, K.M. Chan, E.T. Kool, Fluorescent nucleobases as tools for studying DNA and RNA, *Nat. Chem.*, 9 (2017) 1043-1055.
- [50] D. Baumstark, H.A. Wagenknecht, Fluorescent hydrophobic zippers inside duplex DNA: Interstrand stacking of perylene-3,4 : 9,10-tetracarboxylic acid bisimides as artificial DNA base dyes, *Chem. Eur. J.*, 14 (2008) 6640-6645.
- [51] T. Nguyen, A. Brewer, E. Stulz, Duplex Stabilization and Energy Transfer in Zipper Porphyrin-DNA, *Angew. Chem. Int. Ed.*, 48 (2009) 1974-1977.
- [52] Y.G. Xiang, Q.Y. Zhang, Z.B. Li, H. Chen, Role of electrostatic complementarity between perylenediimide and porphyrin in highly stabilized GNA, *Mater. Sci. Eng. C-Mater. Biol. Appl.*, 70 (2017) 1156-1162.
- [53] E. Meggers, L. Zhang, Synthesis and Properties of the Simplified Nucleic Acid Glycol Nucleic Acid, *Acc. Chem. Res.*, 43 (2010) 1092-1102.
- [54] M. Nakamura, Y. Matsui, T. Takada, K. Yamana, Chromophore Arrays Constructed in the Major Groove of DNA Duplexes Using a Post-Synthetic Strategy, *ChemistrySelect*, 4 (2019) 1525-1529.
- [55] V.T. Dien, M. Holcomb, A.W. Feldman, E.C. Fischer, T.J. Dwyer, F.E. Romesberg, Progress Toward a Semi-Synthetic Organism with an Unrestricted Expanded Genetic Alphabet, *J. Am. Chem. Soc.*, 140 (2018) 16115-16123.
- [56] K. Betz, M. Kimoto, K. Diederichs, I. Hirao, A. Marx, Structural Basis for Expansion of the Genetic Alphabet with an Artificial Nucleobase Pair, *Angew. Chem. Int. Ed.*, 56 (2017) 12000-12003.
- [57] K. Hamashima, M. Kimoto, I. Hirao, Creation of unnatural base pairs for genetic alphabet expansion toward synthetic xenobiology, *Curr. Opin. Chem. Biol.*, 46 (2018) 108-114.
- [58] P. Roethlisberger, A. Istrate, M.J.M. Lopez, R. Visini, A. Stocker, J.L. Reymond, C.J. Leumann, X-ray structure of a lectin-bound DNA duplex containing an unnatural phenanthrenyl pair, *Chem. Commun.*, 52 (2016) 4749-4752.
- [59] J. Ludwig, F. Eckstein, Rapid and efficient synthesis of nucleoside 5'-O-(1-thiotriphosphates), 5'-triphosphates and 2',3'-cyclophosphorothioates using 2-chloro-4H-1,3,2-benzodioxaphosphorin-4-one, *J. Org. Chem.*, 54 (1989) 631-635.
- [60] A. Zahn, C.J. Leumann, Synthesis of functionalized biphenyl-C-nucleosides and their incorporation into oligodeoxynucleotides, *Bioorg. Med. Chem.*, 14 (2006) 6174-6188.
- [61] A. Zahn, C.J. Leumann, Recognition properties of donor- and acceptor-modified biphenyl-DNA, *Chem. Eur. J.*, 14 (2008) 1087-1094.
- [62] P. Röthlisberger, F. Levi-Acobas, I. Sarac, R. Ricoux, J.P. Mahy, P. Herdewijn, P. Marliere, M. Hollenstein, Incorporation of a minimal nucleotide into DNA, *Tetrahedron Lett.*, 59 (2018) 4241-4244.
- [63] T.J. Matray, E.T. Kool, A specific partner for abasic damage in DNA, *Nature*, 399 (1999) 704-708.
- [64] S. Nakano, Y. Uotani, Y. Sato, H. Oka, M. Fujii, N. Sugimoto, Conformational changes of the phenyl and naphthyl isocyanate-DNA adducts during DNA replication and by minor groove binding molecules, *Nucleic Acids Res.*, 41 (2013) 8581-8590.
- [65] E.A. Motea, I. Lee, A.J. Berdis, Insights into the Roles of Desolvation and pi-Electron Interactions during DNA Polymerization, *ChemBioChem*, 14 (2013) 489-498.
- [66] M. Flamme, P. Röthlisberger, F. Levi-Acobas, M. Chawla, R. Oliva, L. Cavallo, G. Gasser, P. Marliere, P. Herdewijn, M. Hollenstein, Enzymatic Formation of an Artificial Base Pair Using a Modified Adenine Nucleoside Triphosphate, *ChemRxiv*, (2019) doi: 10.26434/chemrxiv.11427459.v11427451.
- [67] A. Horhota, K.Y. Zou, J.K. Ichida, B. Yu, L.W. McLaughlin, J.W. Szostak, J.C. Chaput, Kinetic analysis of an efficient DNA-dependent TNA polymerase, *J. Am. Chem. Soc.*, 127 (2005) 7427-7434.
- [68] S. Tabor, C.C. Richardson, Effect of manganese ions on the incorporation of dideoxynucleotides by bacteriophage-t7 DNA-polymerase and escherichia-coli DNA-polymerase-i, *Proc. Natl. Acad. Sci. U. S. A.*, 86 (1989) 4076-4080.
- [69] Y.K. Chang, Y.P. Huang, X.X. Liu, T.P. Ko, Y. Bessho, Y. Kawano, M. Maestre-Reyna, W.J. Wu, M.D. Tsai, Human DNA Polymerase mu Can Use a Noncanonical Mechanism for Multiple Mn²⁺-Mediated Functions, *J. Am. Chem. Soc.*, 141 (2019) 8489-8502.
- [70] S. Vichier-Guerre, L. Dugue, S. Pochet, 2'-Deoxyribonucleoside 5'-triphosphates bearing 4-phenyl and 4-pyrimidinyl imidazoles as DNA polymerase substrates, *Org. Biomol. Chem.*, 17 (2019) 290-301.
- [71] I. Sarac, M. Hollenstein, Terminal Deoxynucleotidyl Transferase in the Synthesis and Modification of Nucleic Acids, *ChemBioChem*, 20 (2019) 860-871.
- [72] L. Tang, L.A. Navarro, A. Chilkoti, S. Zauscher, High-Molecular-Weight Polynucleotides by Transferase-Catalyzed Living Chain-Growth Polycondensation, *Angew. Chem. Int. Ed.*, 56 (2017) 6778-6782.
- [73] M. Hollenstein, F. Wojciechowski, C.J. Leumann, Polymerase incorporation of pyrene-nucleoside triphosphates, *Bioorg. Med. Chem. Lett.*, 22 (2012) 4428-4430.
- [74] Y.J. Cho, E.T. Kool, Enzymatic synthesis of fluorescent oligomers assembled on a DNA backbone, *ChemBioChem*, 7 (2006) 669-672.
- [75] S.V. Vasilyeva, N.A. Kuznetsov, A.S. Kuznetsova, J.G. Khalyavina, D.A. Tropina, T.I. Lavrikova, O.I. Kargina, L.M. Gornostaev, DNA fluorescent labeling with naphtho 1,2,3-cd indol-6(2H)-one for investigation of protein-DNA interactions, *Bioorg. Chem.*, 72 (2017) 268-272.
- [76] A. Johnson, A. Karimi, N.W. Luedtke, Enzymatic Incorporation of a Coumarin-Guanine Base Pair, *Angew. Chem. Int. Ed.*, (2019) 10.1002/anie.201910059.
- [77] Y. Li, A. Fin, L. McCoy, Y. Tor, Polymerase-Mediated Site-Specific Incorporation of a Synthetic Fluorescent Isomorphous G Surrogate into RNA, *Angew. Chem. Int. Ed.*, 56 (2017) 1303-1307.

[78] T.J. Chen, N. Hongdilokkul, Z.X. Liu, R. Adhikary, S.S. Tsuen, F.E. Romesberg, Evolution of thermophilic DNA polymerases for the recognition and amplification of C2'-modified DNA, *Nat. Chem.*, 8 (2016) 557-563.

[79] A. Marx, K. Betz, The Structural Basis for Processing of Unnatural Base Pairs by DNA Polymerases, *Chem. Eur. J.*, 26 (2020) 3446-3463.

1.

Supplementary Material

Supplementary data (additional gel images, NMR and LC-MS characterization of modified oligonucleotides) can be found associated with this article.

SUPPORTING INFORMATION

Enzymatic synthesis of biphenyl-DNA oligonucleotides

Pascal Röthlisberger,^{a,b} Fabienne Levi-Acobas,^{a,b} Christian J. Leumann,^c and Marcel Hollenstein^{a,b*}

^a Institut Pasteur, Department of Structural Biology and Chemistry, Laboratory for Bioorganic Chemistry of Nucleic Acids, CNRS UMR 3523, 28, rue du Docteur Roux, 75724 Paris Cedex 15, France.

^b Institut Pasteur, Department of Genome and Genetics, Paris, France

^c Department of Chemistry and Biochemistry, University of Bern, Freiestrasse 3, CH-3012 Bern, Switzerland

E-mail: marcel.hollenstein@pasteur.fr

SUPPORTING INFORMATION

1. Additional gel images	2-4
2. NMR and MALDI spectra	5-8
3. Oligonucleotide synthesis and characterization	9-16

1. Additional gel images:

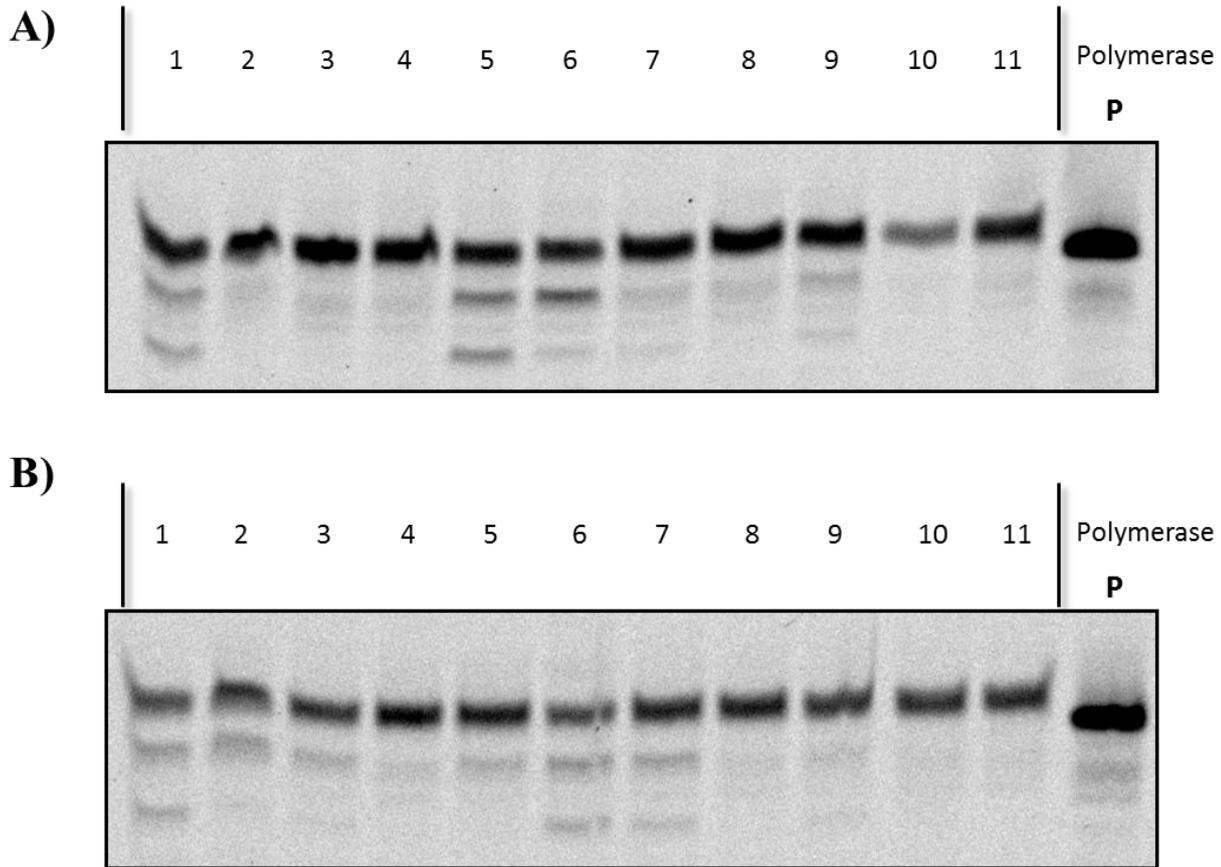


Fig. S1. Gel analysis (PAGE 20%) of PEX reactions with primer **P1**, template **T1**, and **dBphTP** A) and **dφTP** B). All the reaction mixtures contained 200 μ M of nucleotides and 2 U of the DNA polymerases. The reaction mixtures were incubated at the corresponding temperatures in 1X reaction buffer for 2 h. **P** indicates the unreacted primer. Polymerases used: 1. Phusion; 2. Hemo KlenTaq; 3. Q5 DNA polymerase; 4. *Bst*; 5. Taq; 6. Terminator; 7. Vent (*exo*⁻); 8. Dpo4; 9. Deep Vent; 10. Phi29; 11. Kf *exo*⁻.

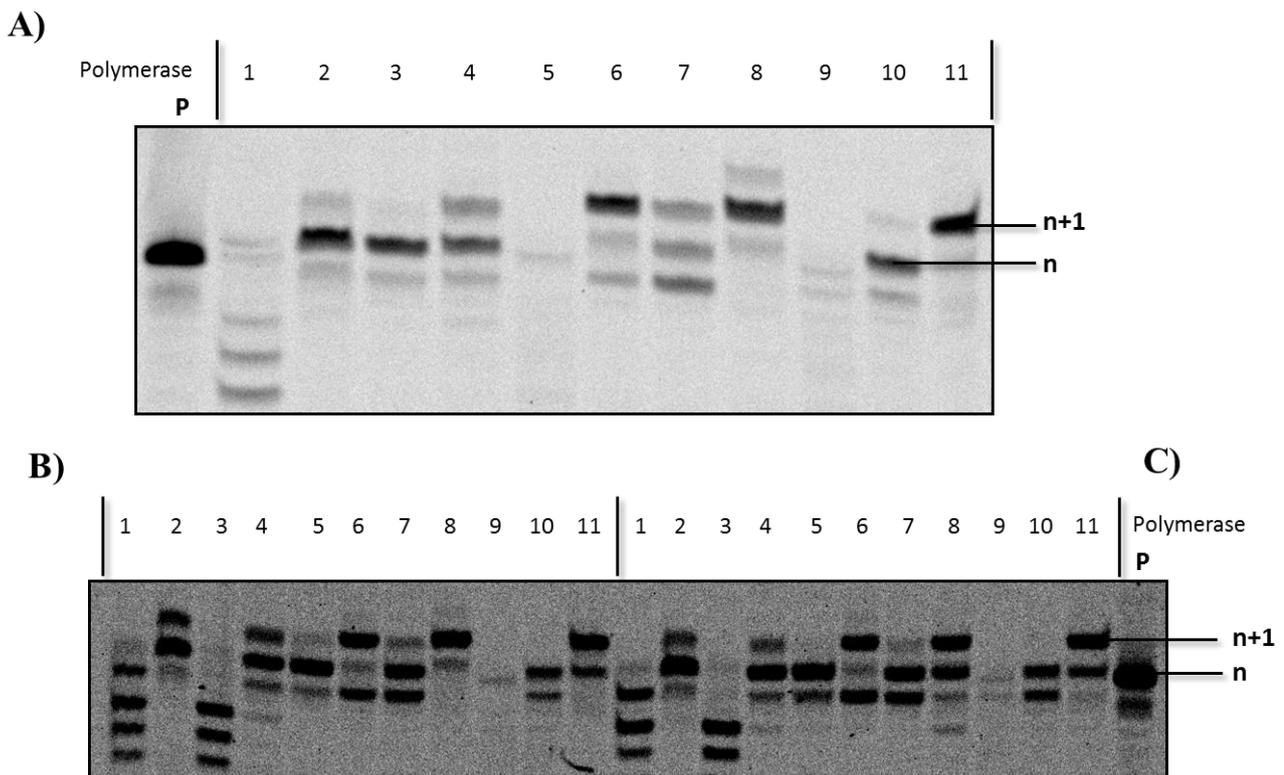


Fig. S2. Gel analysis (PAGE 20%) of PEX reactions with primer **P1**, template **T9** A), **T10** B), and **T11** C) in the presence of **dBphTP**. All the reaction mixtures contained 200 μ M of nucleotides and 2 U of the DNA polymerases. The reaction mixtures were incubated at the corresponding temperatures in 1X reaction buffer for 1 h. **P** indicates the unreacted primer. Polymerases used: 1. Phusion; 2. Hemo KlenTaq; 3. Q5 DNA polymerase; 4. *Bst*; 5. Taq; 6. Terminator; 7. Vent (*exo*⁻); 8. Dpo4; 9. Deep Vent; 10. Phi29; 11. Kf *exo*⁻.

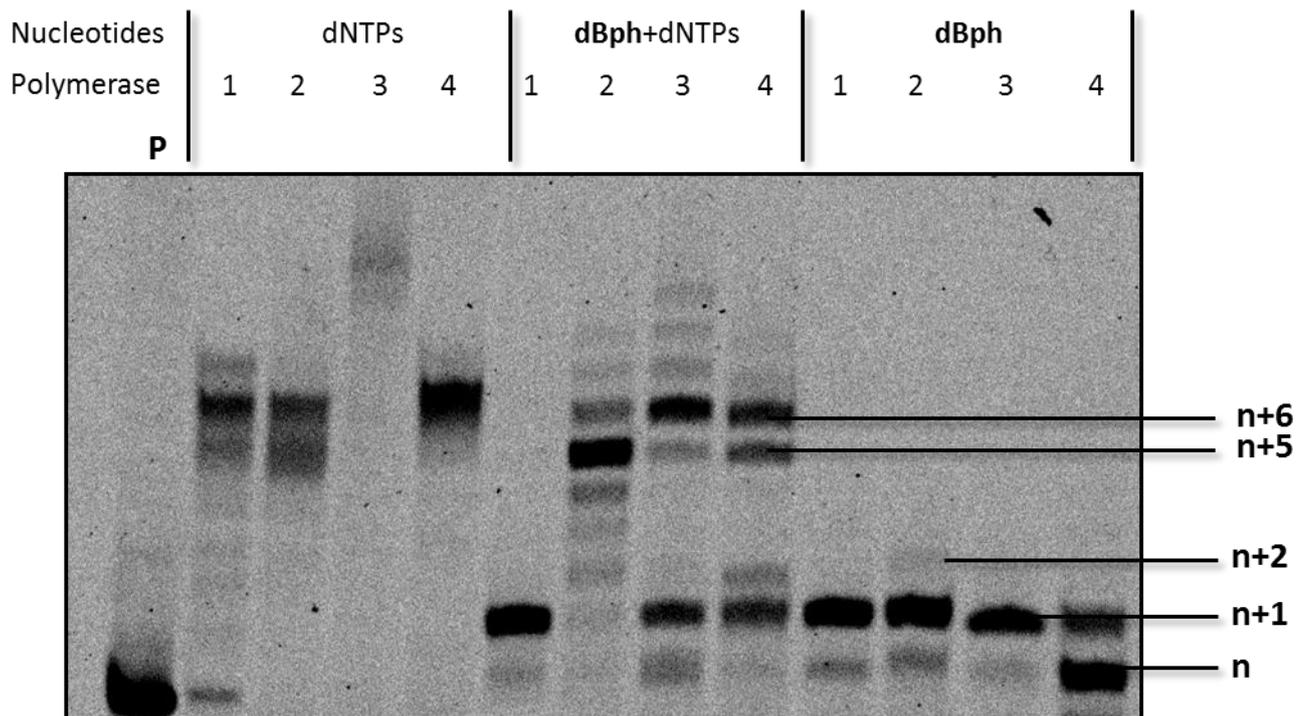


Fig. S3. Gel analysis (PAGE 20%) of products stemming from various bypass experiments. PEX reactions with natural controls were carried out using template **T14**, primer **P1**, and incubate at the adequate temperature for 1 h. The modified **dBph** unit was incorporated opposite a single abasic site (template **T9** as described in Fig. S2) and natural dNTPs were then added and incubated for an additional hour. **P** indicates the unreacted primer. Polymerases used: 1. *Bst*; 2. Terminator; 3. Dpo4; 4. *Kf exo⁻*.

Table S1. Yields (in %) of n+1 product formation based on the gel analysis shown in Figures 4 and 5.

Polymerase	T3	T4	T5	T6	T7	T8
Phusion	0	0	0	0	0	0
Hemo Klen Taq	31	25	15	23	19	19
Q5 DNA polymerase	0	0	0	0	0	0
Taq	14	0	4	5	5	15
<i>Bst</i>	50	20	45	22	18	41
Terminator	76	48	71	42	63	67
Vent (<i>exo⁻</i>)	0	0	0	0	6	0
Dpo4	74	54	62	44	36	71
Deep Vent	0	0	0	0	0	0
Phi29	0	0	0	0	0	21
<i>Kf exo⁻</i>	81	74	67	64	54	71

2. NMR and MALDI spectra

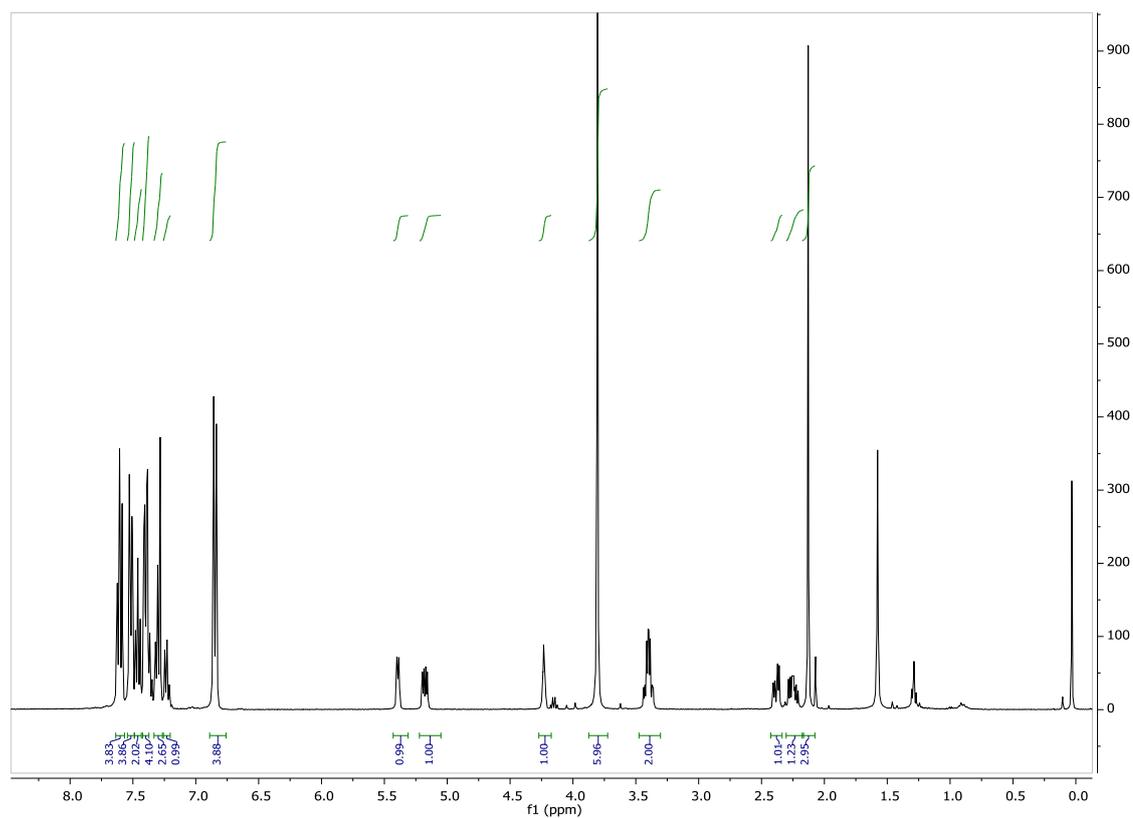


Fig. S4. ^1H NMR spectrum of compound **3**.

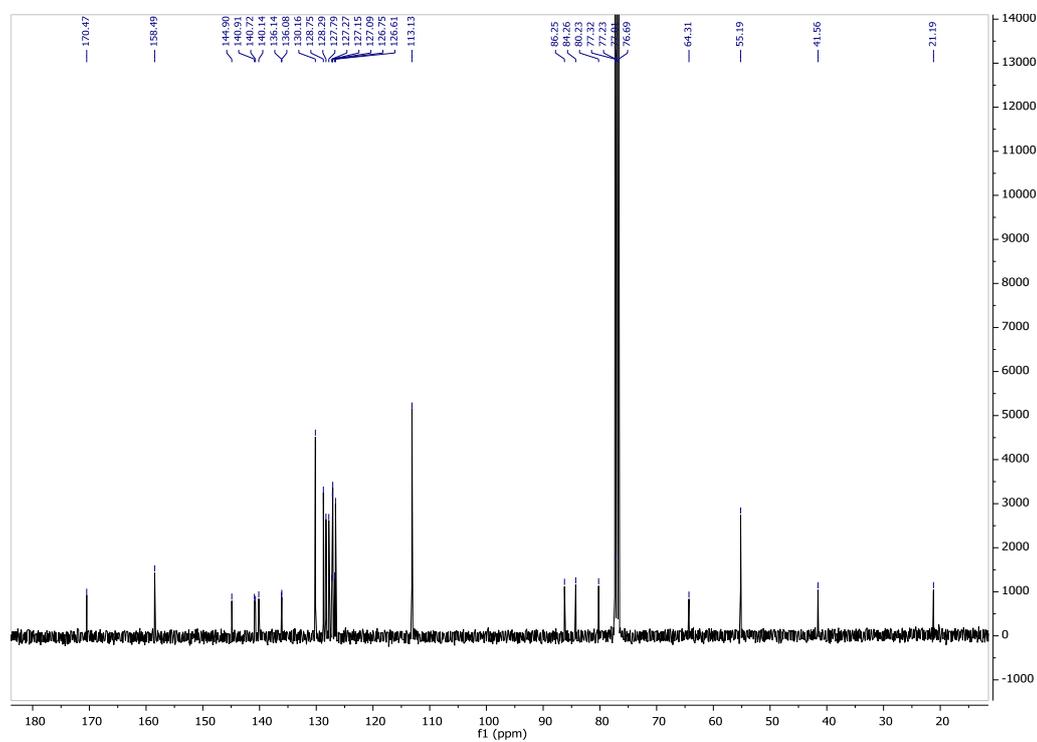


Fig. S5. ^{13}C NMR spectrum of compound **3**.

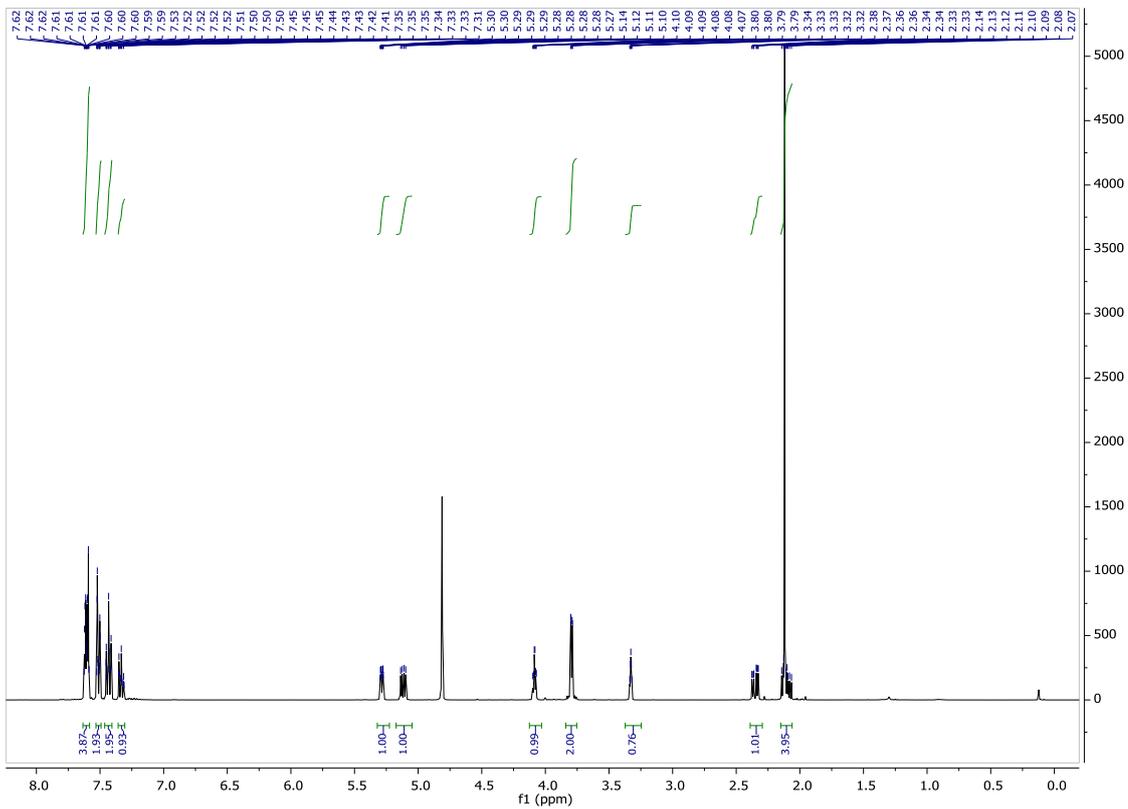


Fig. S6. ^1H NMR spectrum of compound **4**.

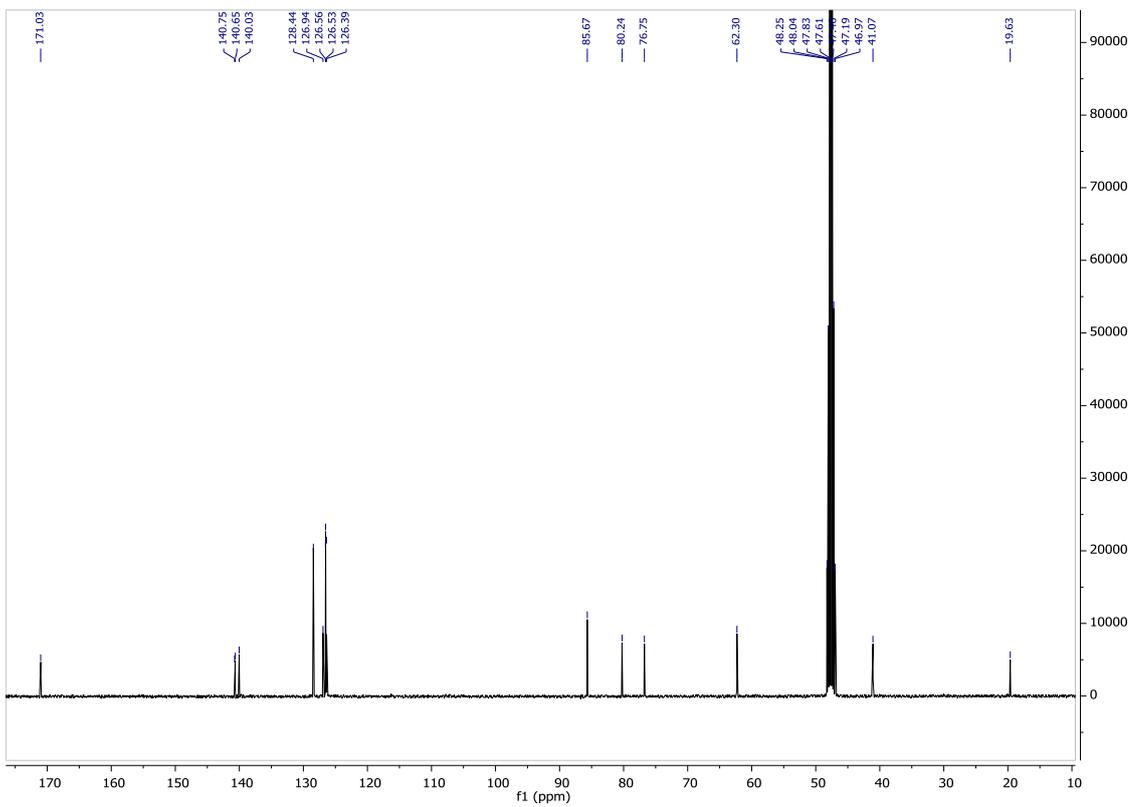


Fig. S7. ^{13}C NMR spectrum of compound **4**.

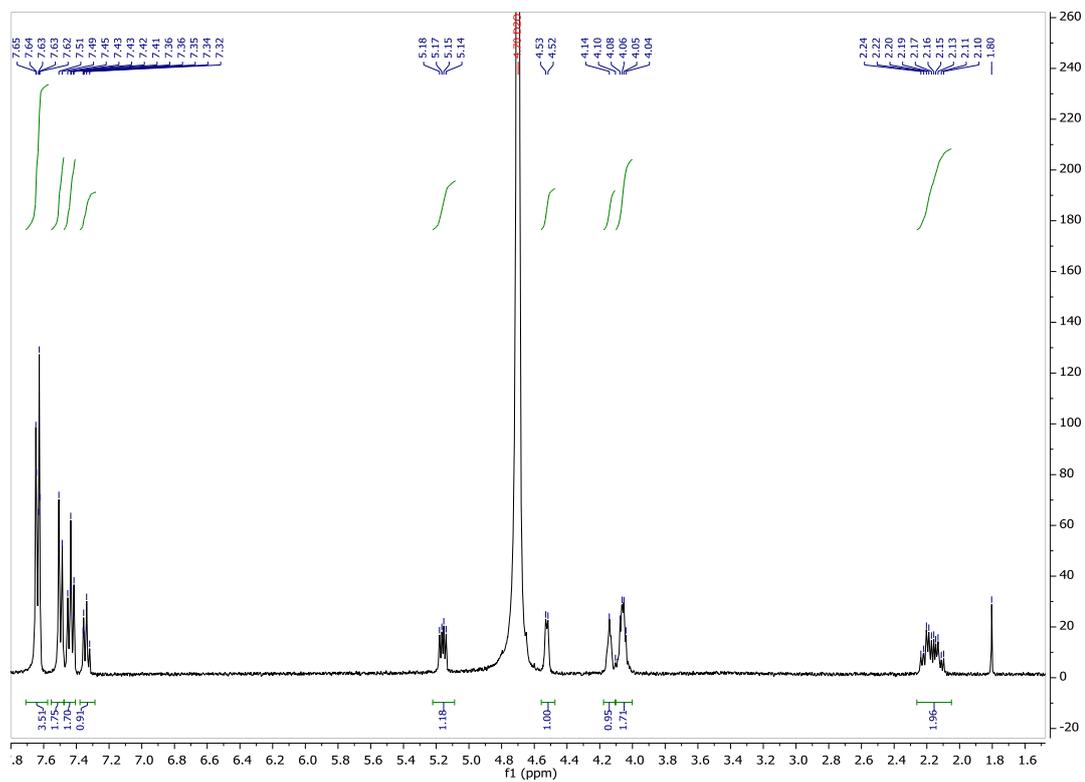


Fig. S8. ^1H NMR spectrum of **dBphTP 1**.

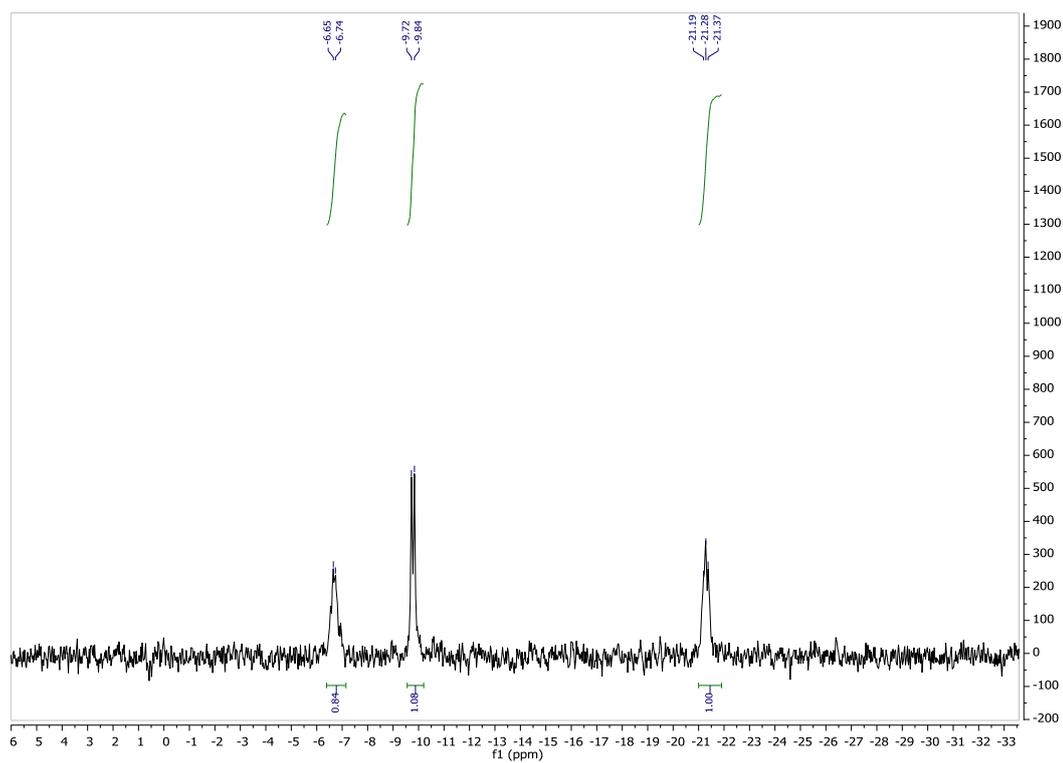


Fig. S9. ^{31}P NMR spectrum of **dBphTP 1**.

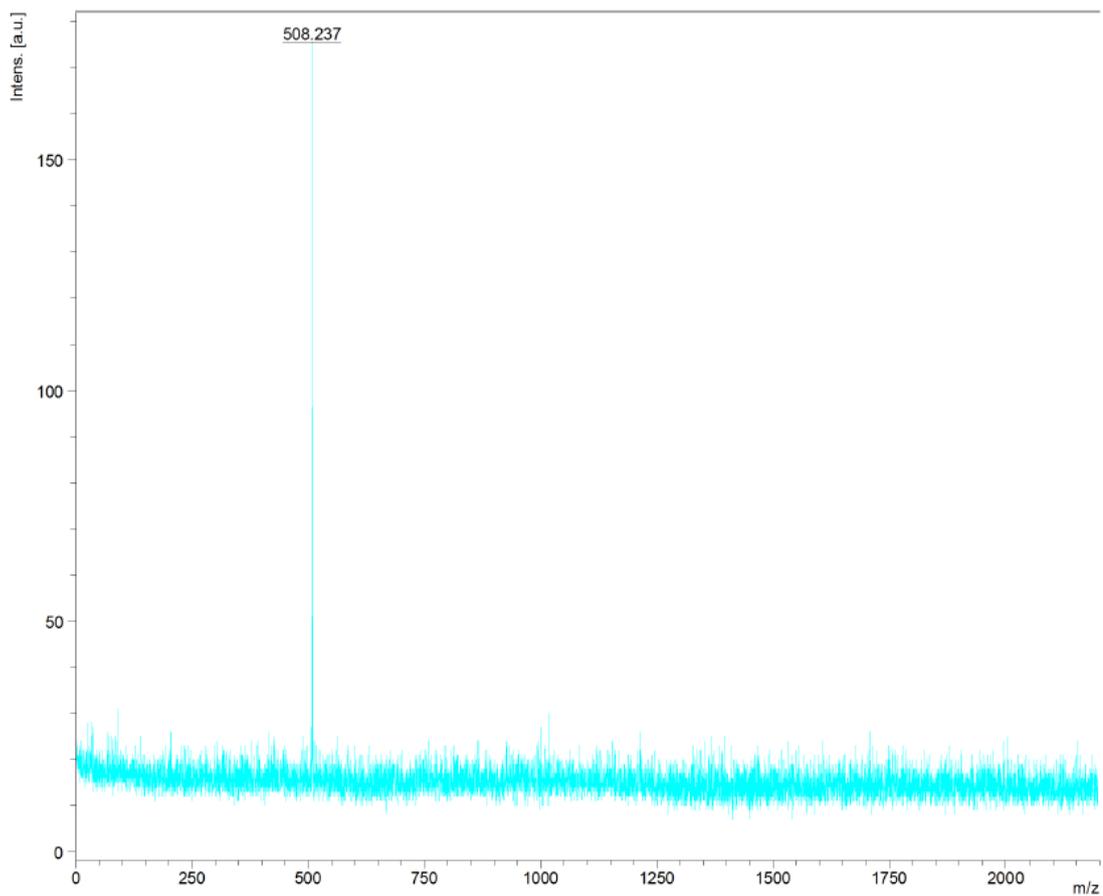


Fig. S10. MALDI MS spectrum of **dBphTP 1**.

3. Oligonucleotide synthesis and characterization:

Table S2. Overview of oligonucleotides containing biphenyl and modified biphenyl nucleotides.

Template	Sequence (shown 5' to 3')	M_{calc.}	M_{obs.}
T1	GGAB ph TGAGGCTATAGTGAGTCGTA	7514	7517
T2	GGB phBphBph GAGGCTATAGTGAGTCGTA	7561	7565
T3	GGAG X1 GAGGCTATAGTGAGTCGTA	7554	7557
T4	GGAG X2 GAGGCTATAGTGAGTCGTA	7554	7555
T5	GGAG X3 GAGGCTATAGTGAGTCGTA	7569	7571
T6	GGAG X4 GAGGCTATAGTGAGTCGTA	7569	7570
T7	GGAG X5 GAGGCTATAGTGAGTCGTA	7584	7586
T8	GGAG X6 GAGGCTATAGTGAGTCGTA	7584	7585
T9	5'-GGA G φ G AGG CTA TAG TGA GTC GTA	7387	7386 ^b
T10	5'-GG φ φ G AGG CTA TAG TGA GTC GTA	7105	7104 ^b
T11	5'- φφφ φφ G AGG CTA TAG TGA GTC GTA	6807	6805 ^b

^a **X1 = 5; X2 = 6; X3 = 7; X4 = 8; X5 = 9; X6 = 10.** ^b data taken from *Tet. Lett.* **2018**, 59, 4241-4244.

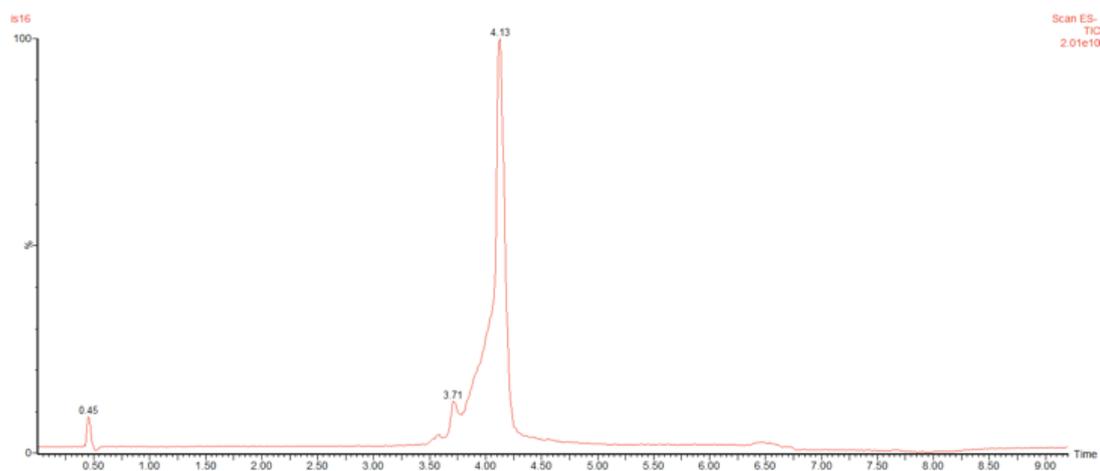


Fig. S11. Total ion chromatogram (TIC) of template **T1**.

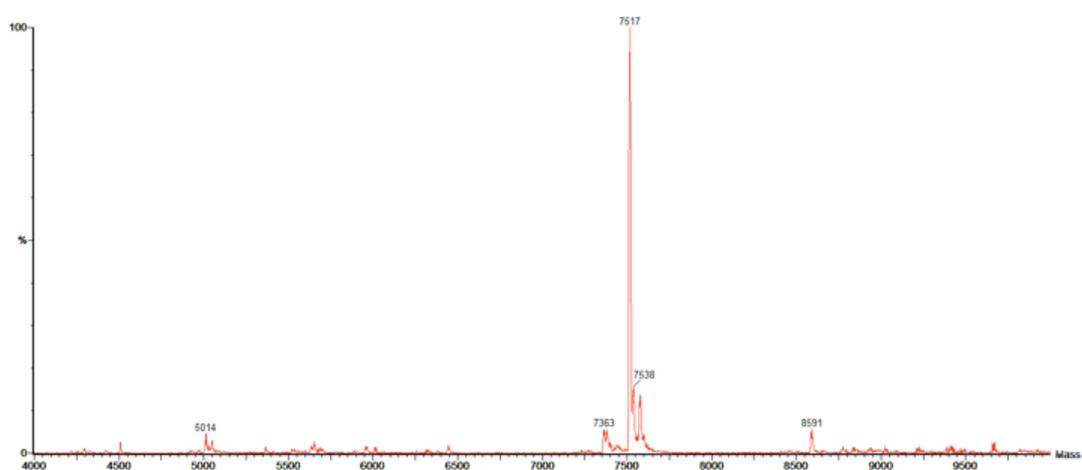


Fig. S12. Mass spectrum of template **T1**.

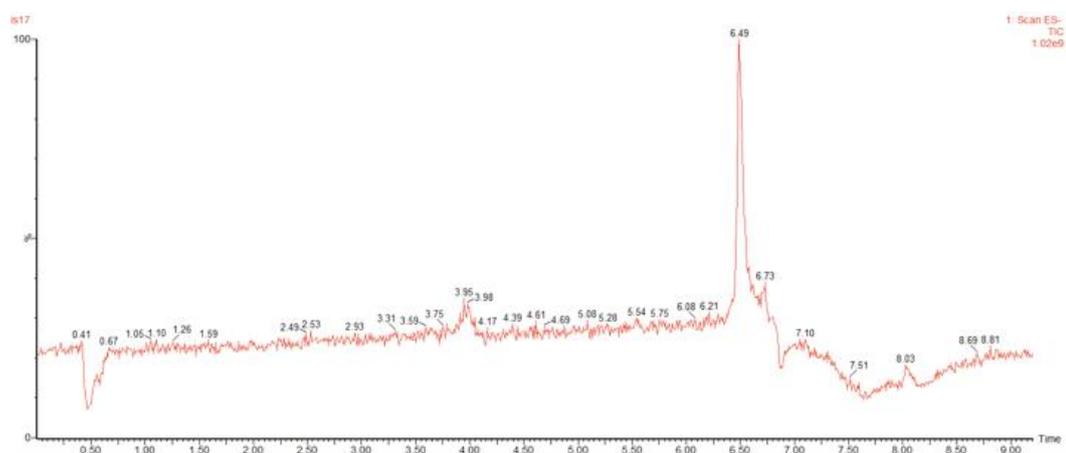


Fig. S13. Total ion chromatogram (TIC) of template **T2**.

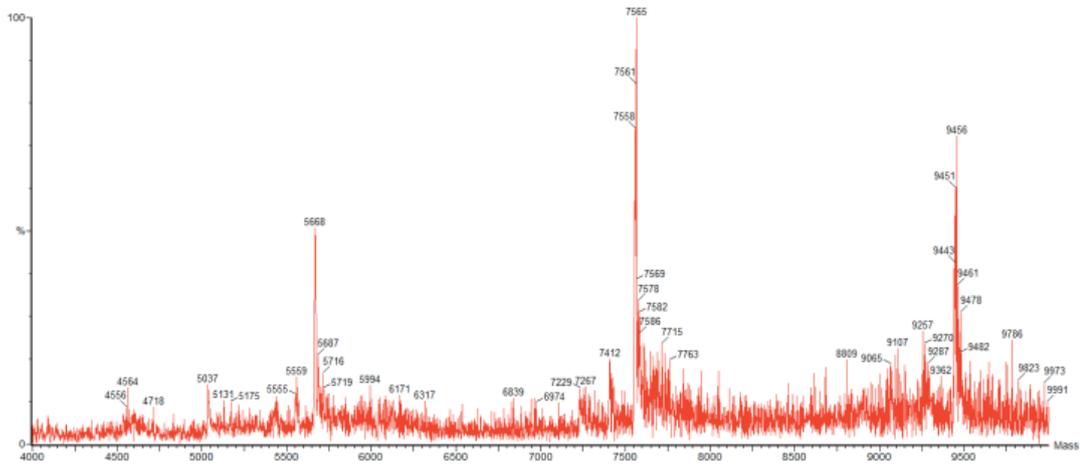


Fig. S14. Mass spectrum of template **T2**.

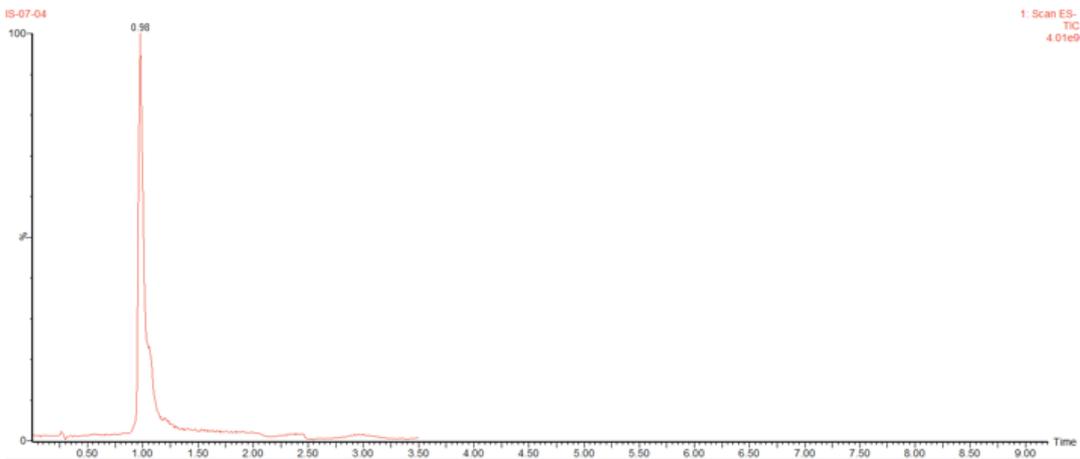


Fig. S15. Total ion chromatogram (TIC) of template **T3**.

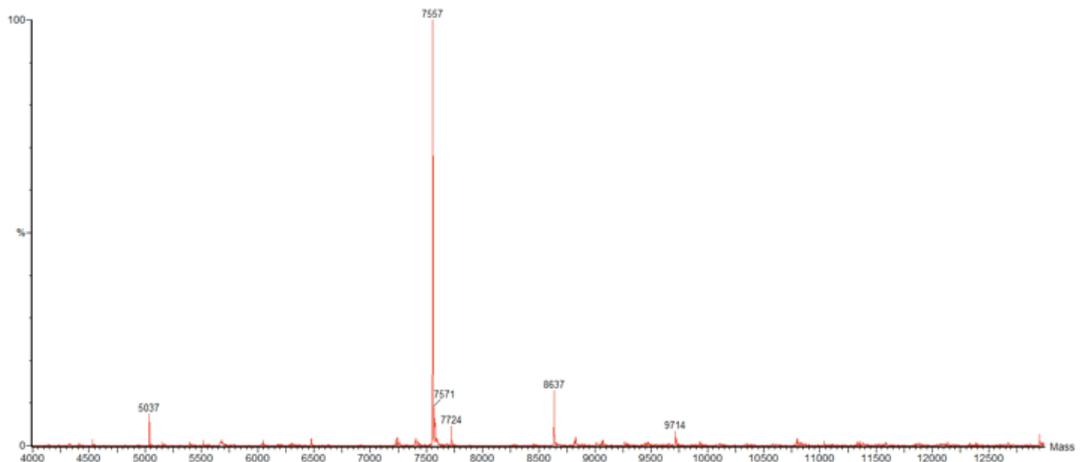


Fig. S16. Mass spectrum of template **T3**.

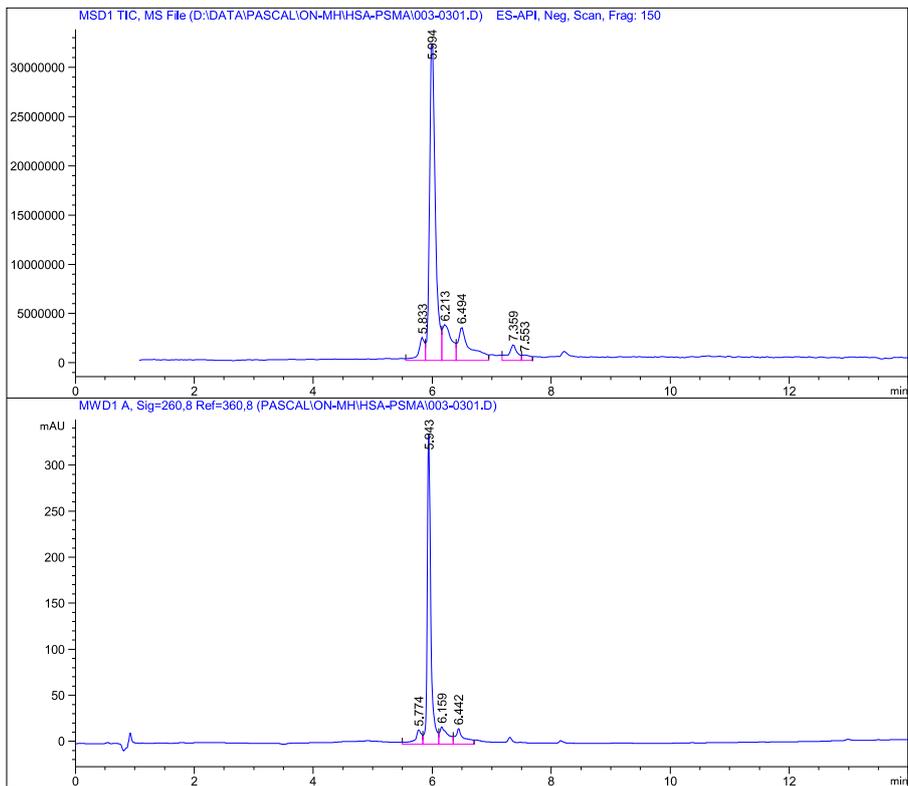


Fig. S17. Total ion chromatogram (TIC) of template **T4**.

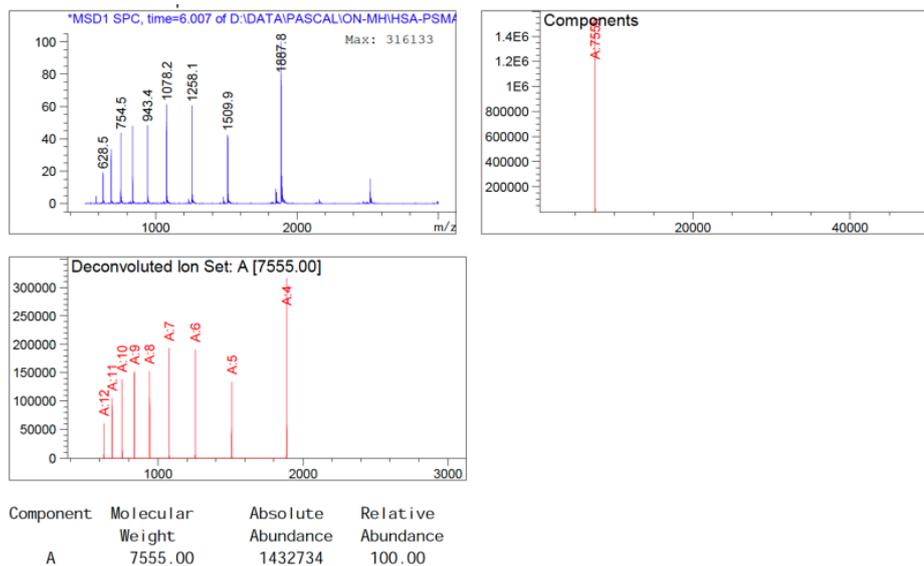


Fig. S18. Mass spectrum of template **T4**.

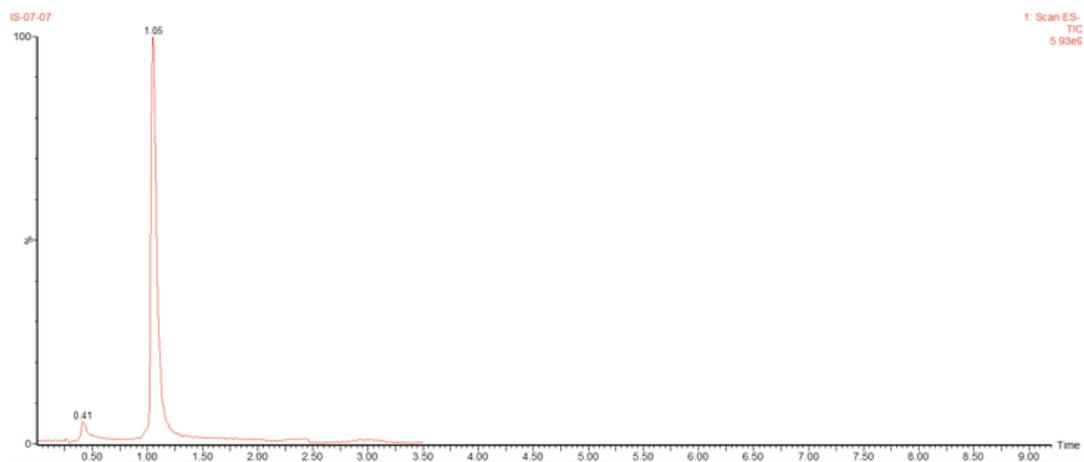


Fig. S19. Total ion chromatogram (TIC) of template **T5**.

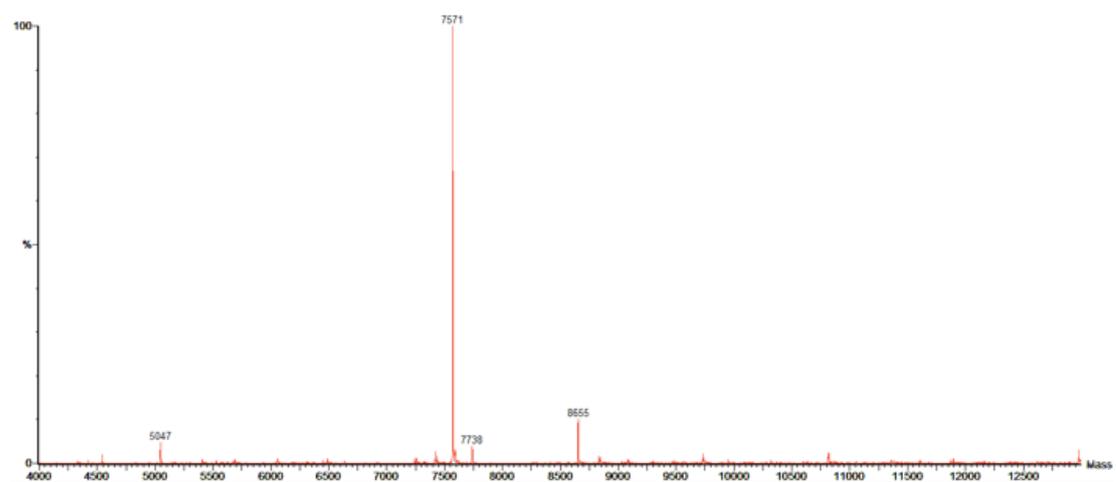


Fig. S20. Mass spectrum of template **T5**.

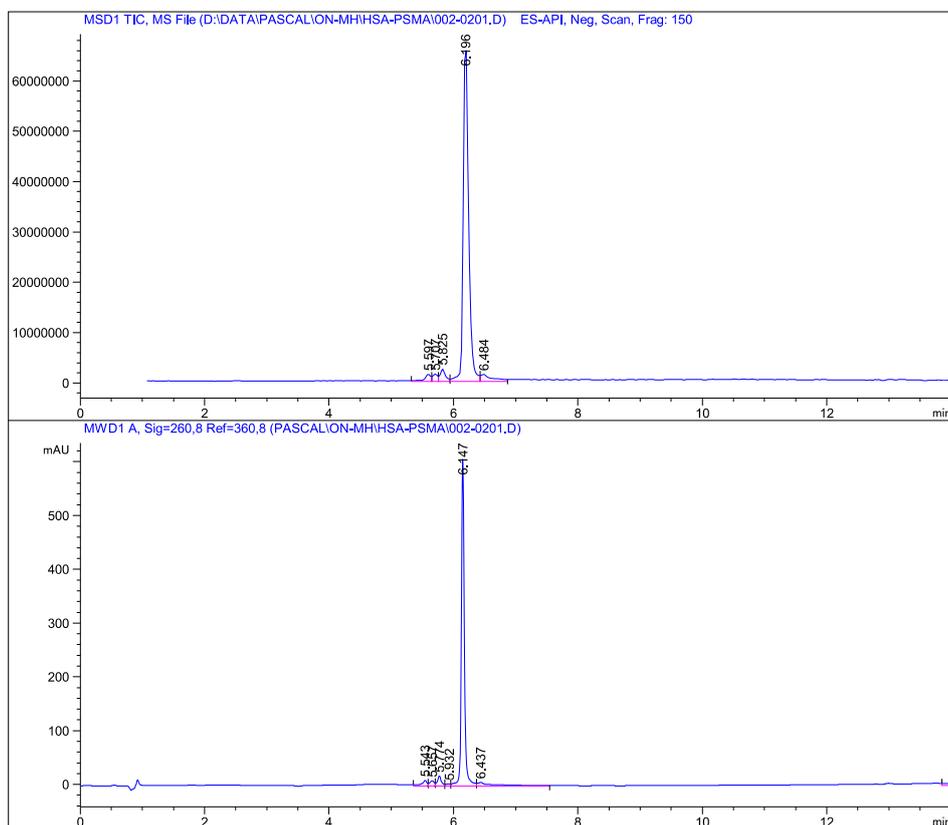


Fig. S21. Total ion chromatogram (TIC) of template **T6**.

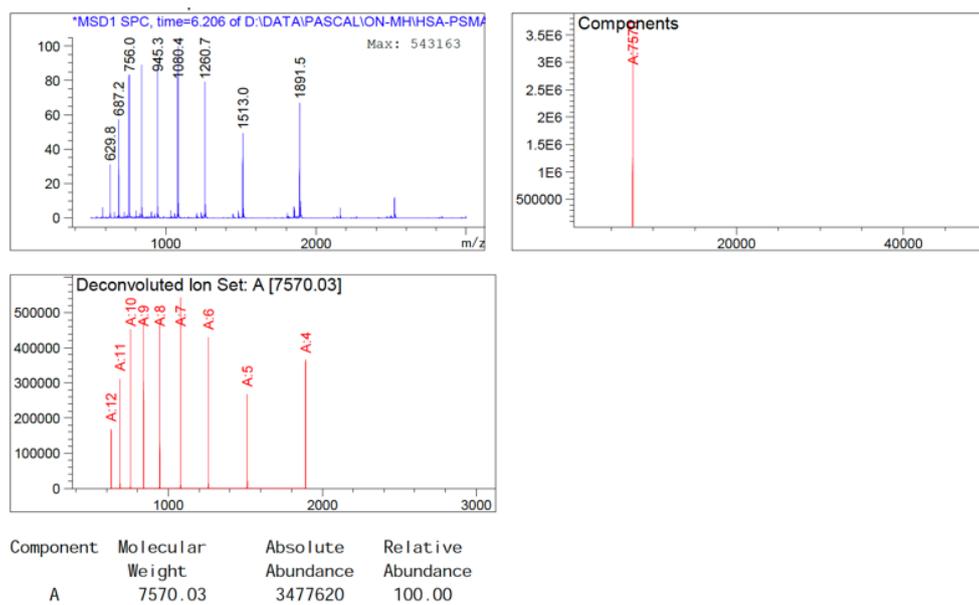


Fig. S22. Mass spectrum of template **T6**.

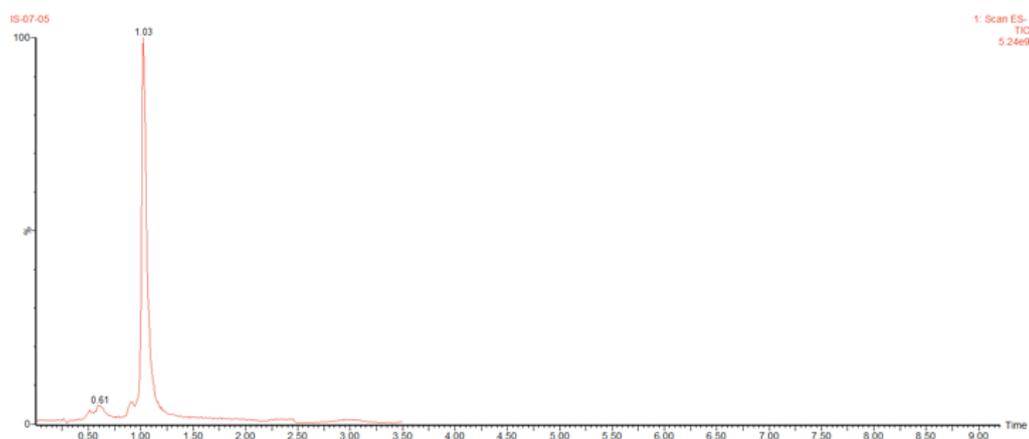


Fig. S23. Total ion chromatogram (TIC) of template **T7**.

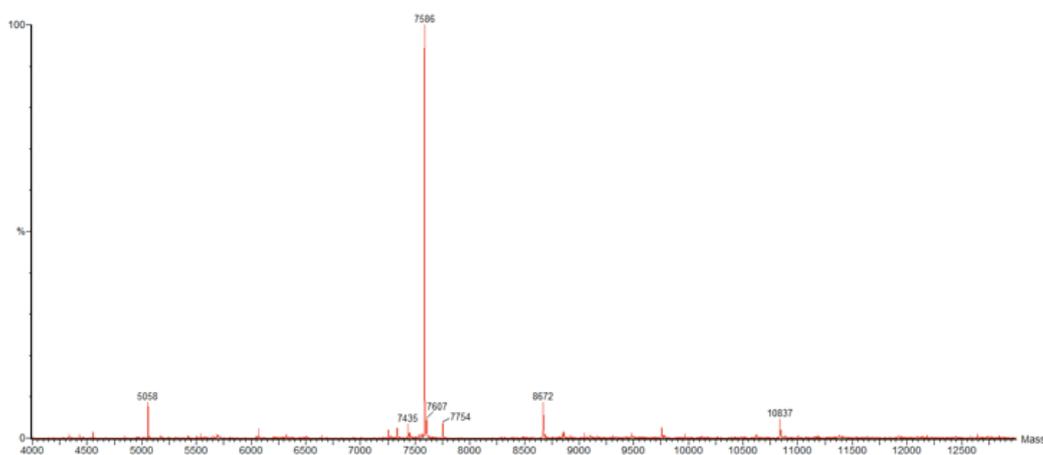


Fig. S24. Mass spectrum of template **T7**.

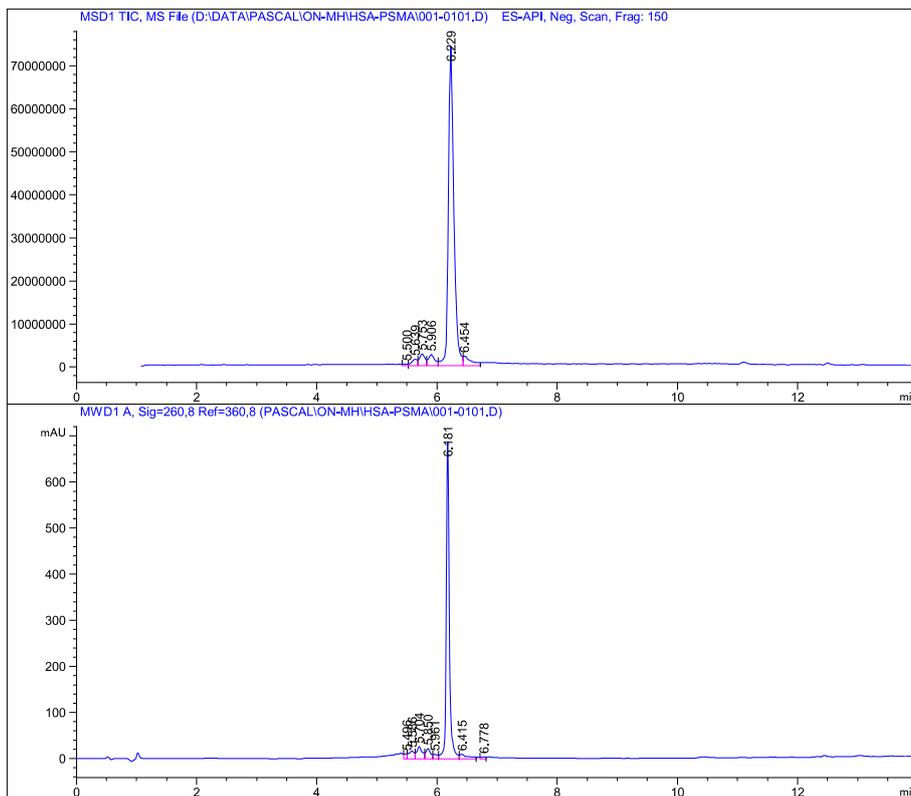


Fig. S25. Total ion chromatogram (TIC) of template **T8**.

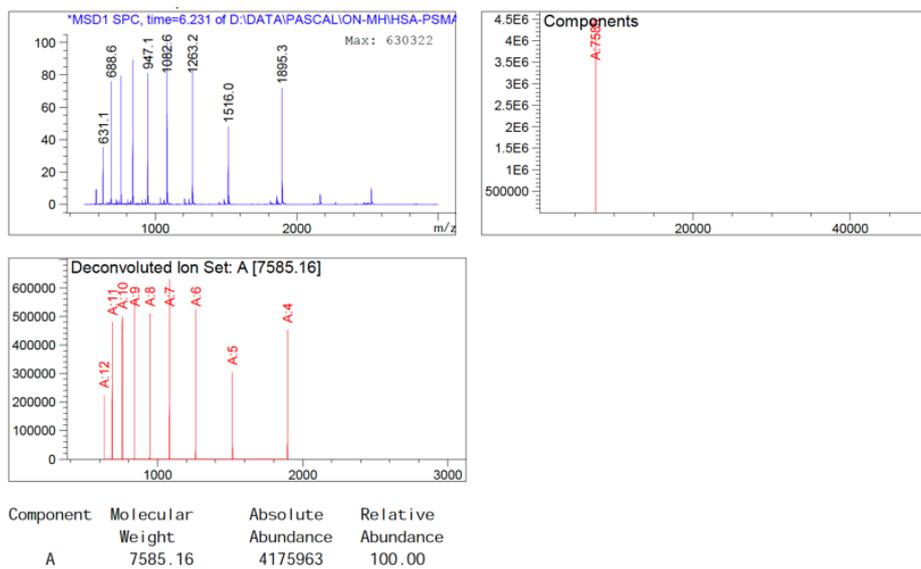


Fig. S26. Mass spectrum of template **T8**.