nature chemistry

SUPPLEMENTARY INFORMATION

https://doi.org/10.1038/s41557-019-0255-4

In the format provided by the authors and unedited.

A synthetic genetic polymer with an uncharged backbone chemistry based on alkyl phosphonate nucleic acids

Sebastian Arangundy-Franklin[1], Alexander I. Taylor[1], Benjamin T. Porebski[1], Vito Genna[2], Sew Peak-Chew[1], Alexandra Vaisman[3], Roger Woodgate[3], Modesto Orozco[2,4] and Philipp Holliger[1]*

[1]MRC Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge Biomedical Campus, Cambridge, UK. [2]Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute of Science and Technology, Barcelona, Spain. [3]Section on DNA Replication, Repair and Mutagenesis, Bethesda, MD, USA. [4]Department of Biochemistry and Biomedicine, University of Barcelona, Barcelona, Spain. *e-mail: ph1@mrc-lmb.cam.ac.uk

Supplementary Materials for

"A synthetic genetic polymer with an uncharged backbone chemistry based on alkyl-phosphonate nucleic acids"

Authors: Sebastian Arangundy-Franklin, Alexander I. Taylor, Benjamin T. Porebski, Vito Genna, Sew Peak-Chew, Alexandra Vaisman, Roger Woodgate, Modesto Orozco & Philipp Holliger*

***Correspondence to:** ph1@mrc-lmb.cam.ac.uk

Supplementary Materials include:

Supplementary Materials and Methods

1. Chemical Synthesis of α-Phosphate Alkylated 2'-deoxynucleoside Triphosphate analogues

The synthesis of phNTPs was carried out using a method analogous to that of Dyatkina[21], with a few modifications. For the synthesis of P-α-ethyl-phosphonate dNTPs we used Ethylphosphonic dichloride as the alkylphosphonylating reagent, whereas Methylphosphonic dichloride was used for the pilot-scale synthesis of P-α-methyl-dUTP and P-α-methyl-dCTP. For the majority of our large-scale synthesis of phNA, we used custom-synthesized phNTPs from Jena (P-α-methyl-dCTP) or GeneACT (P-α-methyl-dUTP and P-α-ethyl-dATP).

1.1 Synthesis of 2'-deoxyadenosine 5'–(α-ethylphosphonyl)-β,γ-diphosphate.

2'-Deoxyadenosine (0.5 mmol) was dissolved in 5 mL triethylphosphate. To this solution were added 1.5 equivalents proton sponge and the solution was cooled to -7°C. 3 Equivalents of ethylphosphonic dichloride ($POCl_2Et$) were added to the solution dropwise. After 5 hours 3.2mL of a 0.5M solution of tributylammonium pyrophosphate in DMF were added along with 300µL of tributylamine. After 5 hours, the reaction was quenched by the addition of 1mL 0.1M TEAB and stored at 4°C. The solution was diluted to 45 mL and loaded onto a DEAE-Sephadex A25 column and eluted with a linear gradient of 0-1M TEAB, fractions containing triphosphate analogue eluted at 0.6M TEAB. The pooled fractions were freeze-dried to a glassy-film before being dissolved in deionized water and purified by ion-paired reversed phase HPLC, eluting as two separate peaks. Pooled fractions were lyophilized, dissolved in water and passed through a DOWEX ion exchange resin to convert it to the sodium salt. The yield was 1.5%. 31P NMR (D_2O) δ = 28.8 (d), 28.2 (d), –6.2 (m), –22.9 (m) ppm. LC-MS: expected m/z = 501.22, measured = 502.

1.2 Synthesis of 2'-deoxycytidine 5'–(α-ethylphosphonyl)-β,γ-diphosphate.

2'-Deoxycytidine (0.5 mmol) was suspended in 5mL triethylphosphate along with 2 equivalents proton sponge, sonicated and heated to 70 °C until a solution formed. To the ice-cooled solution were added 3 equivalents of POCl2Et and the solution was stirred overnight at 4°C before the addition of 3.8 mL of 0.5M tributylammonium pyrophosphate

in DMF and 0.5mL tributylamine. After 16 hours the reaction was quenched with 0.05M TEAB to a final volume of 50 mL, before filtering and purifying by anion exchange as previously described. The purified triphosphate analogue fractions were pooled, lyophilized and dissolved in water prior to further purification by ion-paired RP-HPLC. Unlike other triphosphate analogues synthesized, the dC triphosphate analogue eluted from the column as a single peak in the buffer system used. Pooled elution fractions were lyophilized and converted to the sodium salt, then stored at -20°C. The yield was 1%.31P NMR (D_2O) δ = 28.7 (d), 28.5 (d), –9.5 (m), –23.7(m) ppm. LC-MS: expected m/z = 477.2, measured = 478.0

1.3 Synthesis of 2'-deoxyguanosine 5'–(α-ethylphosphonyl)-β,γ-diphosphate.

2'-Deoxyguanosine (0.5mmol) and 0.165g proton sponge were dissolved in 5mL triethylphosphate by sonication and heating to 70° C. The solution was cooled to 0°C before the dropwise addition of 2 equivalents of $POCl_2Et$, stirred for 2hrs before adding 1.58mmol Tributylammonium pyrophosphate in 3.2mL DMF and 0.5mL tributylamine. After stirring overnight at 4°C the reaction was quenched by the dropwise addition of 50mL 0.05M TEAB. The triphosphate analogue was purified from the reaction mixture by anion exchange chromatography as described above. The relevant fractions were pooled and further purified by RP-HPLC, eluting from the column as a two large, overlapping peaks. The pooled eluate was lyophilized, converted to the sodium salt and stored at -20°C. 31P NMR (D_2O) δ = 28.8 (d), 28.4 (d), –8.2 (m), –23.3 (m) ppm. LC-MS: expected m/z = 517.22, measured = 518.

1.4 Synthesis of Thymidine 5'–(α-ethylphosphonyl)-β,γ-diphosphate.

Thymidine (0.5mmol) and 0.25g Proton Sponge were dissolved in 5mL triethylphosphate. The solution was cooled to 0°C before adding 3 equivalents of $POCl_2Et$ and stirring overnight at 4° C. After nearly complete conversion to the 5' Ethylphosphonic chloride, pyrophosphorylation was carried out by the addition of 10mL of 0.5M Tributylammonium pyrophosphate and 2mL of tributylammonium, with stirring at 0°C. The reaction was quenched with 100mL 0.1M TEAB, and filtered before ion-exchange chromatography (Sephadex A-25). Relevant fractions were pooled and lyophilized to a clear, glassy film. The lyophilized product was dissolved in water before purification by HPLC, which yielded two broadly overlapping peaks, corresponding to each stereoisomer. Pooled fractions

were lyophilized and stored at -20° C. 31P NMR (D_2O) δ = 28.6 (d), 28.2 (d), –7.7 (m), –23.27 (m) ppm. LC-MS: expected m/z = 492.21, measured = 493.

1.5 HPLC for phNTP Diastereomer Separations

Stereoisomers were separated using a Varian Prostar modular HPLC equipped with a Gemni 5µm C6-Phenyl 110A (250x4.6mm) column and an ion-pairing buffer system. Buffer A was 0.1M TEAB pH 7.9, while Buffer B was 0.1M TEAB 90% Methanol pH 7.9. The flow rate for all purifications was 1.5mLmin^{-1} and we diluted the stock nucleotides into buffer A to 2.5mM. For the separation of P-α-Met-dUTP we used a 0-6% Buffer B gradient over 27 minutes. For the separation of P-α-Met-dCTP we used a 0-4% Buffer B gradient over 17 minutes. For the separation of P-α-Et-dATP we used a 5-15% Buffer B gradient over 40 minutes. For the separation of P-α-Et-dGTP we used a 0-10% Buffer B over 30 minutes. All the relevant fractions were pooled and lyophilized prior to polymerase incorporation assays.

2. Polymerase Engineering and phNA characterization

All oligonucleotides used are listed in Supplementary Tables 6-8.

2.1 Preparation of Polymerase Variants

As a starting point for all directed evolution experiments, we used a mutant version of the *Thermococcus gorgonarius* polymerase deficient of exonuclease activity D141A, E143A; deficient of uracil stalling activity V93Q; capable of incorporating various labelled nucleotide analogues A485L and lastly, capable of efficient reverse transcription of several xeno-nucleic acids E429G, I521L, K726R. For simplicity, this polymerase is here referred to as RT521L[2]. For structure-guided mutagenesis, we created saturation mutagenesis mini-libraries, where every residue could be sampled by mutating the wild-type codon with an NNS codon. We used oligonucleotides that would generate the mutated polymerase-carrying plasmid pASK75 through iPCR. For directed evolution experiments, we created several mutant libraries, which diversified clusters of residues in regions of RT521L involved in substrate binding and catalysis. Diversification was carried out by iPCR of the plasmid containing the polymerase gene, using primers encoding semi-conservative, evolutionarily related mutations. Each iPCR reaction contained 1x Expand

High Fidelity buffer, 0.4 µM forward and reverse primers, 30ng template, 200µM dNTPs (Roche), 10 units Expand High Fidelity polymerase (Roche) in 100µL volume. Reactions were carried out using an automated Peltier thermal cycler with the following parameters: 94° C 130'; 25 cycles of 94°C 30", 50°C 30", 68°C 18'. PCR products were purified using the Qiagen PCR purification kit, and eluted in 100uL buffer EB before being digested by BsaI (NEB) in 1X Buffer 4 (NEB), supplemented with 0.1mg/ml Bovine Serum Albumin (BSA) and 20 units BsaI in 120µL volume for 3h at 37°C, 12h at 50C and 20' at 65°C. After silica column purification (Qiagen), digested library DNA was ligated for 2 hours at room temperature in a reaction containing 800 units T4 ligase in 1x T4 Ligase Buffer (NEB) for 2 h at room temperature. The DNA was then purified by extraction with phenol : chloroform : isoamyl alcohol (25:24:1) pH 8 and precipitated with isopropanol. The resulting pellet was dissolved in buffer EB, dried by speed-vacuum to remove excess alcohol. The concentrated ligated DNA was used for the transformation of either chemically competent or electro competent *E. coli* 10-β cells (NEB) according to the manufacturer's instructions and plated in TY agar supplemented with 1% glucose. Colonies carrying individual mutants were grown in 200µL wells in 2xTY growth medium, then induced for polymerase expression for 4 h by the addition of anhydrotetracycline. Cells were transferred to a PCR plate, harvested by centrifugation, and resuspended in 30µL 1xThermopol Buffer (NEB) before lysis at 80°C for 30 min. Lysates were cleared by centrifugation at 4000g before polymerase activity assays.

2.2 Polymerase Screening by Polymerase Activity assay (PAA)

PAA screening was carried out essentially as described[3]. Typical primer extensions used 5µL lysate, 1µM Biotinylated primer bFD (5'-Biotin-TEG-CCCCTTATTAGCGTTTGCCA-3'), 5µM template TempN (5'-CTCACGATGCTGGACCAGATAAGCACTTAGCCACGT AGTGCTGTTCGGTAATCGATCTGGCAAACGCTAATAAGGGG-3'), 250 µM phNTPs, 1xThermopol Buffer. Primer extension products were bound to a Streptavidin-Coated plate (Roche) wells directly after the enzymatic reaction to at 5nM in PBS buffer supplemented with 0.1% v/v Tween20 (PBST) for 30 min. Templates were removed by chemical denaturation with two washes of 0.1M Sodium Hydroxide (NaOH). The plate was then neutralized with PBST before the addition of a Digoxigenin-labelled oligonucleotide probe, which hybridizes to the primer extension products with increasing

strength depending on the length of primer extension. After washing away excess probe, an anti-DIG Horseradish Peroxidase antibody fragment (Roche) was allowed to bind the probe in the plate wells. The assay was then developed by the addition of 1-Step™ ULTRA TMB-ELISA substrate (Thermo), stopped with 1M Sulfuric acid (H_2SO_4), and absorbance is measured using a Spectramax 340PC384 plate reader (Molecular Devices), measuring absorbance at 450nm.

2.3 Polymerase Directed Evolution by Compartmentalized Self-Tagging (CST)

Cells expressing polymerase variants were prepared as described above. Approximately $2x10^8$ cells were used to prepare the aqueous emulsion, containing 1xThermopol Buffer, 20mM $MgCl_2$, 10µM-100nM Biotinylated primer, 10% v/v glycerol, 2% v/v Formamide, 0.1mg/mL BSA, 133µM phNTPs in the aqueous phase totalling 150µL; 7% v/v ABIL WE09, 20% v/v Mineral Oil, 73% v/v Tegosoft DEC in the oil phase, totalling 600µL. The 750µL emulsion was split into four tubes and primer extensions are carried out in a thermal cycler using the following temperature incubations:

Cycle number		1	2	3	4	5
Extension time (min)	4.25	94				
	1	94	94	94	94	94
	15	37	37	50	60	65
	15	50	65	65	65	65
		Temperature (˚C)				

After thermal cycling, 50µL buffer EB were added to the emulsion. The emulsion was broken by extraction with saturated hexanol, once with 1mL hexanol and once with 700 µL hexanol. After removing the organic phase, the aqueous phase was passed through an S400 gel filtration column (GE Healthcare) in order to remove excess biotinylated primers not bound to plasmids. The filtrate was then added to 10µL MyOne C1 beads (Life Technologies) in a 500µL volume of BWBS buffer (10mM Tris HCl pH 7.5, 1mM EDTA, 1M NaCl, 0.1% v/v TWEEN20) and plasmid capture was allowed to take place overnight on an overhead rotator. In order to remove plasmids bound non-specifically to beads, the beads were washed as follows: 500µL BWBS, 500µL TBT2 (10 mM Tris HCl

pH 7.5, 20 mM NaCl, 0.1 mg/ml BSA, 0.1% TWEEN 20), 20% Formamide, 500µL EBT (1x Buffer EB, 0.1% TWEEN). Beads are eluted in 100µL EBT and stored at 4˚C until amplification and PCR.

Amplification of active polymerase genes was carried out by PCR using primers that allow for the amplification of the polymerase domain of RT521L, excluding the N-terminal exonuclease domain. PCR reactions were carried out using 5µL magnetic beads from selection, 0.2µM TgoBa 578 primer, 0.2 µM TgoFo 308 primer, 0.2mM dNTPs, 1x Expand PCR Buffer, 5 Units Expand Polymerase mix. Thermal cycling was performed with the following parameters: 94˚C 1min 30s; 30 cycles of 94˚C 30s, 50˚C 30s, 68˚C 3min; a final extension step of 72˚C for 5min. The other segment of the plasmid was amplified in a similar manner, using the primers TgoFo 574 and TgoBa 304. After amplification, PCR products were purified using the Qiagen PCR clean-up kit and eluted in buffer EB; digestion of PCR products was carried out using 40 Units BsaI, 1x Buffer 4, 0.1mg/mL BSA, in a 120µL volume. Digest reactions were incubated for 4 h at 37˚C, 12 h at 50˚C and 30 min at 65˚C, and then purified using the Qiagen PCR clean-up kit, prior to purification by Agarose gel electrophoresis and Agarose gel extraction using silica columns (Qiagen). Ligation of the purified digested PCR products was carried out in 100µL, typically using 320ng insert (1.8kb fragment), 200ng Vector backbone (3.4kb fragment), 1xT4 Ligase buffer, 200 Units T4 Ligase, at room temperature for 2 h. Ligated DNA is subsequently purified by phenol:chloroform extraction as above, precipitated with isopropanol, and used to transform NEB10β cells, in a manner similar to that of the starting library transformation.

2.4 Mass Spectrometry of polymerase-synthesized phNA

We synthesized phNA from a DNA primer oligonucleotide "FD" (5'-CCCCTTA-TTAGCGAAAGCCA-3') using a template encoding for 25 phNA incorporations TSD (5'-ATGCTGTATCGCGTATGCATCGCATCTGGCAAACGCTAATAAG-GGG-3'). Three different products of the primer extension were PAGE purified on a denaturing 8M Urea gel, excised from the gel, electroeluted, isopropanol-precipitated and desalted by dialysis using a 2kDa cut-off membrane (Thermo). Oligo samples, 0.6 µl in water, were spotted onto a MALDI target followed by 0.6 µl of 3-hydroxypicolinic acid. Some oligo samples were vacuum dried, resuspended in 25 µl, 0.1 M TEAA (triethylammonium acetate) and

desalted using a zip-tip C18 (Merck Millipore, USA). The zip-tip C18 was washed 3× 10 µl of 0.1 M TEAA and then 3× 10 µl water. Next, the oligo was eluted directly onto a MALDI target with 5 µl of 3-hydroxypicolinic acid. All mass spectrometric measurements were carried out in positive ion mode on an Ultraflex III TOF-TOF instrument (Bruker Daltonik).

2.5 MALDI-ISD experiments

30µl - 70 µl of the given samples were concentrated in a Speed Vac (Savant) to less than 10ul and centrifuged to settle some pink precipitates. The supernatant were removed and 1 µl of 50% MeCN was added to re-solubilise the precipitates. 0.6 µl of each sample was spotted onto a MALDI target followed by 0.6 µl of 3-hydroxypicolinic acid and air-dried. The dried sample spots were washed with 1 µl of water, air-dried and further 0.6 µl of matrix was added. All ISD experiments were carried out in positive ion mode on an Ultraflex III TOF-TOF instrument (Bruker Daltonik,). The spectra were processed using Flex Analysis v3.3 (Bruker Daltonik) and annotated manually.

2.6 Reverse Transcription of phNA

The scheme is summarized in Supplementary Fig. 8. phNA was transcribed from DNA template TnD (5'-CCCTCCTTCTTCCTCTTCCCCTCACGATGCTGCGACTCAGATA CAGCACTTAGCCACGTAGTGCTGTTCGGTACATCGATCTGGCAAACGCTAATCAGG GGCTATTATTAATT-3') using the 6-FAM labelled NAPfd DNA primer (5' -FAM-CAGTATCGACAAAGGACCCCTTATTAGCGTTTGCCA-3'). Transcription products were PAGE-purified and isopropanol-precipitated before polydeoxyadenylation using 6G12 polymerase (1µgmL^{-1}) and 1mM dATP for 2hr at 65°C. Poly(dA)-phNA oligomers were purified by isopropanol precipitation, then reverse-transcribed using the biotinylated bL3polyT primer (5'- Biotin-TEG-CAGGAAACAGCTATGACAAAT$_{(24)}$ -3'), RT521L polymerase (0.5µgmL^{-1}), 1mM dNTPs at 50°C for 1 hr, 55°C 1hr, 60°C 1hr and 65°C 1hr. The RT products were bound to M1 Streptavidin-coated magnetic beads (Invitrogen), washed with bead-washing buffer (10mM Tris-HCl, 200mM NaCl, 0.1% Tween-20) and the washed beads were preserved in elution buffer (10mM Tris-HCl, 0.1%Tween-20). Beads were used directly in PCR reactions using the NAP and L3 primers (5'-CAGTATCGACAAAGGA-3', 5'- CAGGAAACAGCTATGACAA-3'). Semi-outnesting PCR was carried out using the NAP/Ts7 primer pair (NAP and Ts7 5'-

CCCTCCTTCTTCCTCTTCCC-3'). PCR products were cloned using the TOPO-TA cloning kit (Invitrogen) and cloned products were transformed into 10β cells (NEB), and individual colonies were sequenced by Sanger sequencing.

2.7 Selection of phNA aptamers against streptavidin

Library synthesis was started by the extension of a DNA hairpin, which serves as both primer and template. Hairpin extension for the first round of selection was carried out using 1nmol Hairpin library (HPN25 or HPB40), in a reaction containing 1μM HP, 1xThermopol Buffer, 0.25mM phNTPs, 4mM $MgSO_4$,1μg/mL PGV2 polymerase; the reaction mix was split into ten 500μL microcentrifuge tubes and incubated in a thermal cycler for 5 cycles of 1.5 h at 50°C, 15 min at 65°C and 5 min at 4°C. After cycling, HP extension products were purified using the Qiagen Nucleotide Removal Kit (other protocols such as phenol : chloroform extraction and isopropanol precipitation lead to loss of product due to co-precipitation with protein aggregates) and eluted in buffer EB. A trimming step was carried out to remove any unextended hairpins using Exonuclease III (NEB) in 1xNEB Buffer 1, incubated for 1 h at 37°C, after which Exo III was inactivated by incubation at 70°C for 30 min. A 3-fold excess of the hairpin display oligonucleotide AntiHPTag was annealed to the hairpin library by incubating the mix at 94°C for 5 min, then slowly cooling it to 4°C at 0.1°C/s. The AntiHPTag oligonucleotide was then extended against the DNA template using 1unit of Bst2.0 polymerase (NEB) in the same solution as the Exo III digest, supplemented with 1xThermopol and in the presence of 0.2mM dNTPs; the reaction mix was then incubated in a thermal cycler for 30 min at 65°C. After the strand-displacement reaction, the phNA displayed on the DNA double-stranded duplex was purified by silica-column as above and eluted in buffer EB. Unextended primers (or primers extended non-specifically by the Bst 2.0 polymerase) were digested by incubation with 20 Units Exonuclease I in 1xExo I buffer for 1 h at 37°C and the protein is inactivated by incubation for 1 h at 65°C.

The library was incubated with 1mg MyOne C1 Streptavidin beads (Life Technologies) for the first round of selection, in 1x Binding buffer (1xBWB, 0.8x Thermopol buffer) in a 1mL total volume for 1 h with rotation at room temperature. The binding solution was then split into five aliquots, mixed with 300μL BWBT buffer and washed three times with 500μL

BWBT using the Kingfisher Robot, and each set of beads eluted in 100μL EBT. Bound aptamers were first amplified by PCR using the Test7 primer, which targets the duplex DNA, and the AntiHPOutnest primer, which is used to amplify selectively only molecules for which the strand displacement reaction had been successful. PCR was carried out in a 50μL volume, using 2.5 Units Fastart Taq polymerase, 1xFastart Buffer, 1% Formamide v/v using 20μL beads as template. Thermal cycling used 95°C for 30s denaturation, then 40 cycles of 95°C for 15s, 50°C for 20s, 68°C for 20s, and a final extension step of 72°C for 60s. PCR products were then concentrated using a 30kDa spin filter, and purified by agarose gel electrophoresis using 4% low-melt agarose. PCR products are extracted from the cut agarose slices using the Qiagen gel extraction kit and eluted in buffer EB.

In order to regenerate the starting Hairpin library, the first amplification products were used as template in a second PCR reaction, using the L3Test7 primer (to add a fluorescent label) and a BsaIFDminus tag (used to avoid PCR artifacts). The PCR conditions were identical, except 10ng of the in-nest PCR are used as template in a 5mL total volume PCR (split into 100μL reactions) and the annealing temperature during cycling was 61°C. The PCR reactions were pooled and concentrated and desalted using a Millipore centrifuge filter unit with a 50kDa cut-off before being digested in a 5mL reaction containing 1x Cutsmart Buffer (NEB), 2000 Units BsaI-HF (NEB) and incubated at 37°C overnight. Digested PCR products were desalted by spin-filtration as above, and digested in a 1mL reaction containing 1x Lambda Exonuclease buffer (NEB) and 650 units Lambda exonuclease for 2 h at 37°C. The single-stranded PCR products were purified by phenol-chloroform extraction and isopropanol precipitation, then loaded onto a 10% polyacrylamide 8M-urea gel and run for 2 h at 400V. The bands corresponding to the BsaI digested PCR product were excised and the PCR products allowed to diffuse out of the gel into 800μL buffer EB after breaking the gel using a steal bead and the TissueLyser (Qiagen) at 30rpm for 2 min. The gel pieces were filtered out using a spin filter and the PCR products were isopropanol-precipitated, pelleted and after washing with 70% isopropanol, dissolved in buffer EB and the residual isopropanol removed by speed-vacuum. The single stranded PCR products could now be annealed to form a hairpin and start the next round of library synthesis and selection. For the second round of selection the ratio of target to library was decreased to 5, then 2 and 1 for the third and fourth rounds

of selection. For the third round of selection, the beads were incubated in 1mM biotin following the wash steps, and only molecules that were eluted from the beads during the biotin incubation were taken forward in the selection. For the fifth round of selection, the ratio of target to library was 1:8, but the library was washed as in the previous round, and then subjected to three 5-hour washes before elution in buffer EBT.

2.8 ELONA streptavidin binding assays

Selected and control phNA aptamer sequences were synthesized from a FITC-labelled DNA primer (or 2'O-methyl primer) and a DNA template and PAGE-purified as described previously. For binding assays the DNA portion of the phNA aptamer was hybridized to its complementary sequence at a 1:1 ratio. All sequences were hybridized and assayed in phosphate-buffered saline (PBS) and 0.1% Tween. Aptamer clones were bound to Streptavidin wells (or Neutravidin, or Goat IgG controls) for 1 hour, then unbound aptamers were washed three times with buffer, and an anti-FITC Fab-HRP (Roche) was bound at 0.15 units mL^{-1}, and unbound Fab was removed. After washing the wells three times with wash buffer, the assay was developed using Ultra-TMB ELISA reagent (Thermo). Peroxidase reactions were stopped with 1M H_2SO_4 before absorbance measurements (450 nm).

2.9 Surface Plasmon Resonance with phNA aptamers

Surface Plasmon Resonance (SPR) measurements were made using a BIAcore 2000 instrument (GE Healthcare UK Ltd., UK) at a flow rate of 20 µl min^{-1} in running buffer (140 mM NaCl, 2.5 mM KCl, 1.6 mM KH_2PO_4, 15 mM Na_2HPO_4, 0.02% Tween-20, pH 7.5) at 20°C. For all experiments, immobilized streptavidin (S4762-.5MG, Sigma) or neutravidin (Pierce) surfaces (~4000 or 20,000 RU per flow cell respectively) were prepared using an amine coupling kit (GE Life Sciences UK Ltd, UK) with CM4 sensor chips (GE Life Sciences UK Ltd, UK) in 5 mM sodium acetate, pH 5.5, then equilibrated overnight in running buffer. DNA or phNA aptamers in running buffer were injected for 150 s at a series of concentrations (500 nM, 250 nM, 125 nM, 62.5 nM and 31.3 nM) and dissociation monitored for 1,200 s. Binding data was double-referenced by subtracting data from a blank flow cell as well as molar equivalent injections of "R0". Data were fitted to a parallel association with biphasic dissociation model by non-linear least squares as described in equation 1 and 2.

Equation 1: Parallel association

$$R = R_{eq1}\left(1 - e^{-(K_{a1}Cn_1 + K_{d1})(t-t0)}\right) + R_{eq2}\left(1 - e^{-(K_{a2}Cn_2 + K_{d2})(t-t0)}\right)$$

Where, t is time in seconds, k_{ai} is the association rate constant for component I, R_{eqi} is the steady state response level for component i, C is the molar concentration of the analyte, n_i is the steric interference factor (set as 1), t0 is the start time for the association, k_{di} is the dissociation rate constant for component i.

Equation 2: Biphasic dissociation

$$R = R_1 e^{-k_{d1}(t-t0)} + (R_0 - R_1)e^{-k_{d2}(t-t0)}$$

Where t is time in seconds, R_1 is the initial response level for component 1, k_{di} is the dissociation rate constant for component i, R_0 is the total response level at the start of dissociation, t0 is the start time for the dissociation.

Deviation from pseudo-first order kinetics for streptavidin (binding to biotin) is well established with ligand heterogeneity arising from cooperativity[4], and likely compounded by aptamer heterogeneity or conformational flexibility. Curves fit for phNA aptamers T1-20, T5-20 (Supplementary Fig. 10).

2.10 Steady-state kinetics of nucleotide incorporation.

The general sequence of DNA templates used for the steady-state kinetic experiments was 5'-TCGATACTGGTACTAATGATTAACGAA**XY**AAGCACGTCCGTACCAT, where XY is either TA, TG, TC, or AT. The 17-mer primer (5'-ATGGTACGGACGTGCTT) was ^{32}P-5'-end-labeled and annealed to the DNA templates at a 1:1.5 molar ratio. The steady-state kinetic parameters K_m and V_{max} were measured as a function of nucleotides concentration. DNA templates (10 nM) were incubated with DNA polymerases (20 pM) at 37°C in reaction mixtures containing 20 mM Tris-HCl, pH 7.6, 10 mM $MgCl_2$, 10 mM dithiothreitol, 250 µg/ml bovine serum albumin, 2.5% glycerol, and variable concentrations of nucleotides. Less than 20% of the primers were extended under these conditions. Nucleotide concentrations ranged from 0.1 to 20 µM for dNTPs and from 2.5 µM to 1 mM for phNTPs. Initial time-course studies indicated that accumulation of reaction product was linear for up to 4 min. Therefore, reactions were performed for 2 min and

terminated by mixing with 1 volume of formamide loading dye containing 500 mM EDTA, 0.1% bromophenol blue in 90% formamide and were then immediately transferred onto ice. Products were resolved by denaturing polyacrylamide gel electrophoresis (8M urea, 18% acrylamide, 1.5h at 90W) and then visualized and quantified using a Fuji FLA-5100 Phosphorimager and ImageGauge software. The velocity of nucleotide incorporation was determined as described previously[5, 6]. V_{max} (the maximum reaction velocity) and K_m (nucleotide concentration at which the reaction velocity is half-maximal) were determined from a Hanes-Woolf plot by linear least squares fit. Values for the apparent k_{cat} were calculated based on the assumption that polymerases were fully active and obtained by dividing V_{max} (in nM primer extended per minute) by the enzyme concentration (0.02 nM). The efficiency of nucleotide insertion by pol was calculated as V_{max}/K_m or k_{cat} /K_m. The results are the mean (±S.E.) of four to seven independent determinations.

2.11 Thermal stability Measurements

We measured the thermal stability of DNA duplex and the chimeric phNA-DNA as well as the DNA primer used to synthesize the phNA-DNA chimera to its complementary DNA sequence. Briefly, we mixed equimolar amounts of each strand at 0.1µM in Hybridization buffer (10mM HEPES, 200mM NaCl, 1mM EDTA). Each mix was annealed by incubating at 95°C 3 min then cooling to 4°C at a rate of 0.1° C/sec. The “plus” strand was labelled with a Fluorescein (6-FAM, IDT) moiety at its 5’-end while the “minus” strand was labelled with a BlackHole Quencher 1 (IDT) at its 3’-end. The annealed mix was aliquoted to 4 separate qPCR reaction tubes and the fluorescence intensity was measured using a RotorGene-6000 (Qiagen). Fluorescence intensity during denaturation was measured from 25-95°C while hybridization fluorescence intensity was measured from 95-25° C in steps of 0.5°C with an incubation time of 5 sec per step. Apparent T_m were determined by fitting the normalized fluorescence data to a sigmoidal (variable slope curve) using Prism7 (GraphPad) and the first derivative was determined similarly, smoothing over 4 neighbors with a 2nd order smoothing function.

3. Molecular dynamics (MD)

The six investigated systems were generated starting from the RB69 DNA Pol/dNTP/dsDNA ternary complex crystallographically caught (PDBid 4M3R) at 2.07 Å resolution[7]. DNA primer strand was modified adding methyl/ethyl groups on the non-bridging *pro-S_p* oxygen of the phospho-backbone. Thus, the P-alkyl-phosphonodiester backbone contains mixture of both P-α-Met-phNTP and P-α-Et-phNTP nucleotides.

3.1 Modelling: P-α-Met-phATP and P-α-Et-phATP incoming nucleotides in the active site of RB69 DNA Pol/DNA binary complex were investigated with two different alkyl substitutions and two different conformations. Precisely: *i)* with the alkylic group (Methyl (Met) / Ethyl (Et)) replacing the *pro-S_p* oxygen of P-α-phATP (namely Met-S_p and Et-S_p); *ii)* with the alkylic group replacing the *pro-R_p* oxygen of P-α-phNTP nucleotides (namely, Met-R_p and Et-R_p); iii) and with an imposed rotation along the P-α-phNTP ε-dihedral angle on systems Met-R_p and Et-R_p in order to enhance exploration of the P-α-phNTP conformational space (namely, Met-R_pr and Et-R_pr).

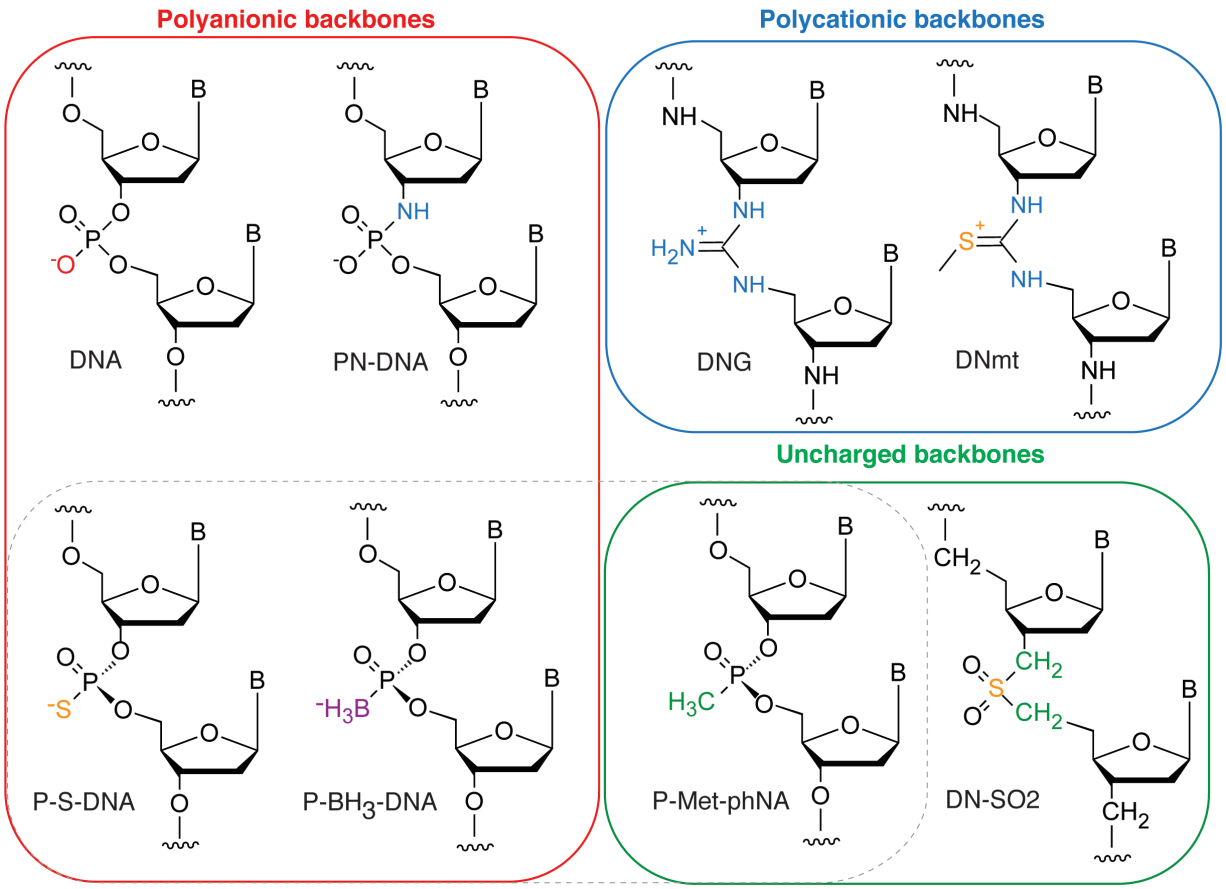
3.2 Force Field-based MD simulations: P-α-Met-phATP and P-α-Et-phATP incoming nucleotides were treated with the general amber force-field (GAFF)[8]. The atomic charges were derived by fitting the electrostatic potential (with B3LYP/6-31G* level of theory) according to Merz-Singh-Kollman scheme[9], also known as the RESP fitting procedure. The length of all covalent bonds, including hydrogen atoms, was set using the LINCS algorithm, allowing a time-integration step of 2 fs. All simulations were performed using Gromacs 2018 code[10]. Long-range electrostatic interactions were calculated with the particle mesh Ewald method with a real space cut-off of 12Å. Periodic boundary conditions in the three directions of Cartesian space were applied. Constant temperature (310 K) was imposed using Langevin dynamics with a damping coefficient of 1 ps[11]. A constant pressure of 1 atm was maintained with Langevin-Piston dynamics with a 200 fs decay period and a 50 fs time constant. The metal active site was treated with a flexible non-bonded approach based on the so called “atoms in molecules” partitioning scheme of the DFT-BLYP electronic density of the active site[12], an approach widely used for similar systems[13, 14, 15]. We could thus consider the charge-transfer interactions between Mg^{2+} ions and their ligands, permitting possible structural rearrangements at the active site

during the MD simulations. All the simulated systems were hydrated using TIP3P water molecules[16]. Na^+ and Cl^- ions were added to neutralize the total charge. The size of the final systems was approximately 133 × 125 × 130 Å, with ~42 000 water molecules, resulting in a total number of ~130 000 atoms each.

3.3 MD simulation protocol: We adopted the following simulation protocol: the systems were minimized using a steepest-descent algorithm and then slowly heated up to 310 K in 10 ns for a total of 2000 steps. The first 50 ns of production run are considered as the equilibration phase. ~1 μs of MD simulations has been collected in the *NPT ensemble* for each of the six investigated systems, resulting in a total of ~6 μs of dynamics. Coordinates of the systems were collected every 5 ps, for a total of ~200 000 frames for each run. Statistics were collected considering the equilibrated trajectories only, thus discarding the first ~50 ns of simulation for all the systems.

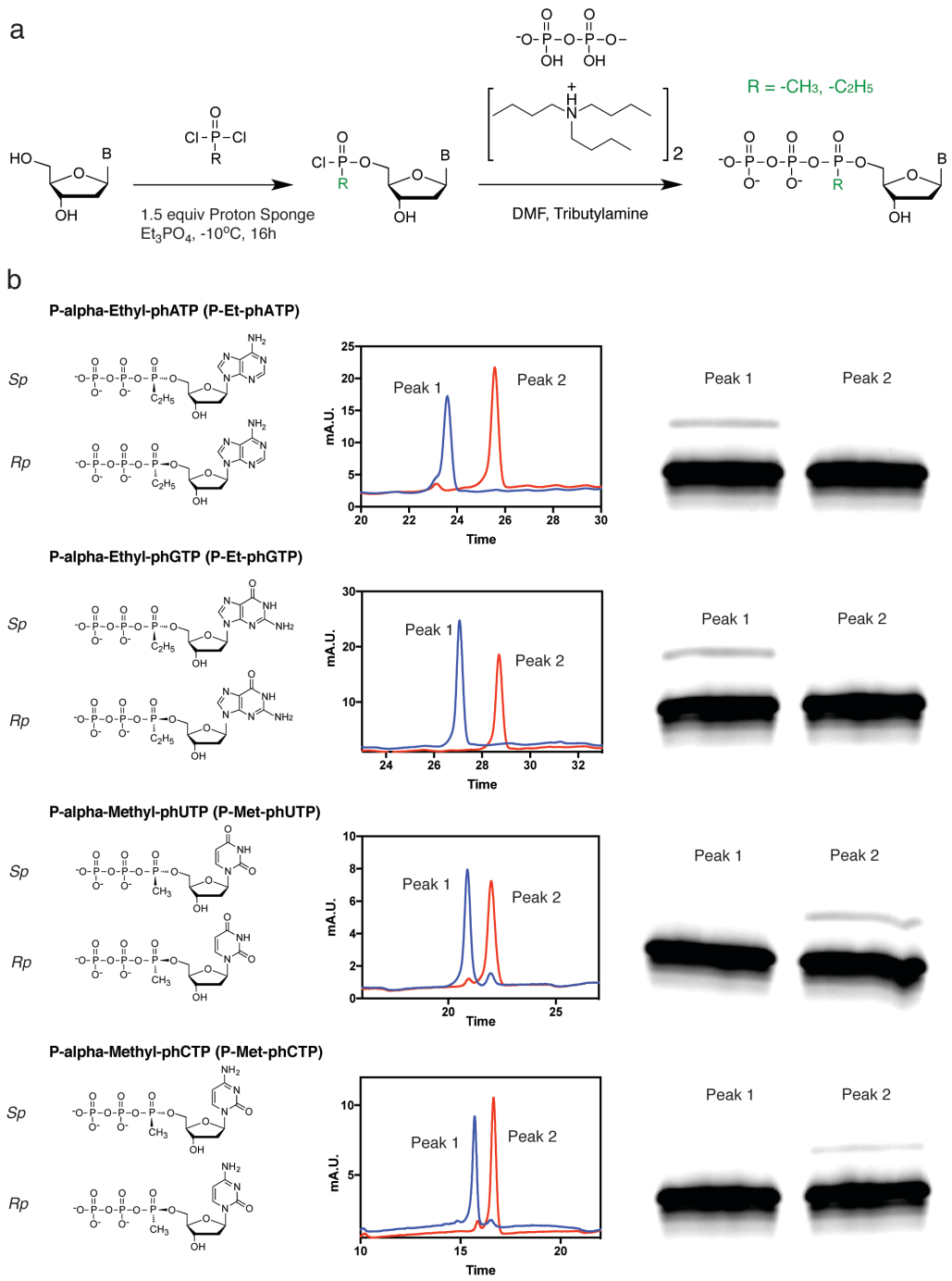
3.4 Analysis of the trajectories: bond lengths, RMSD and Watson-Crick pattern analysis were performed with VMD 1.9.1 package and in-house software[17].

Supplementary Figures



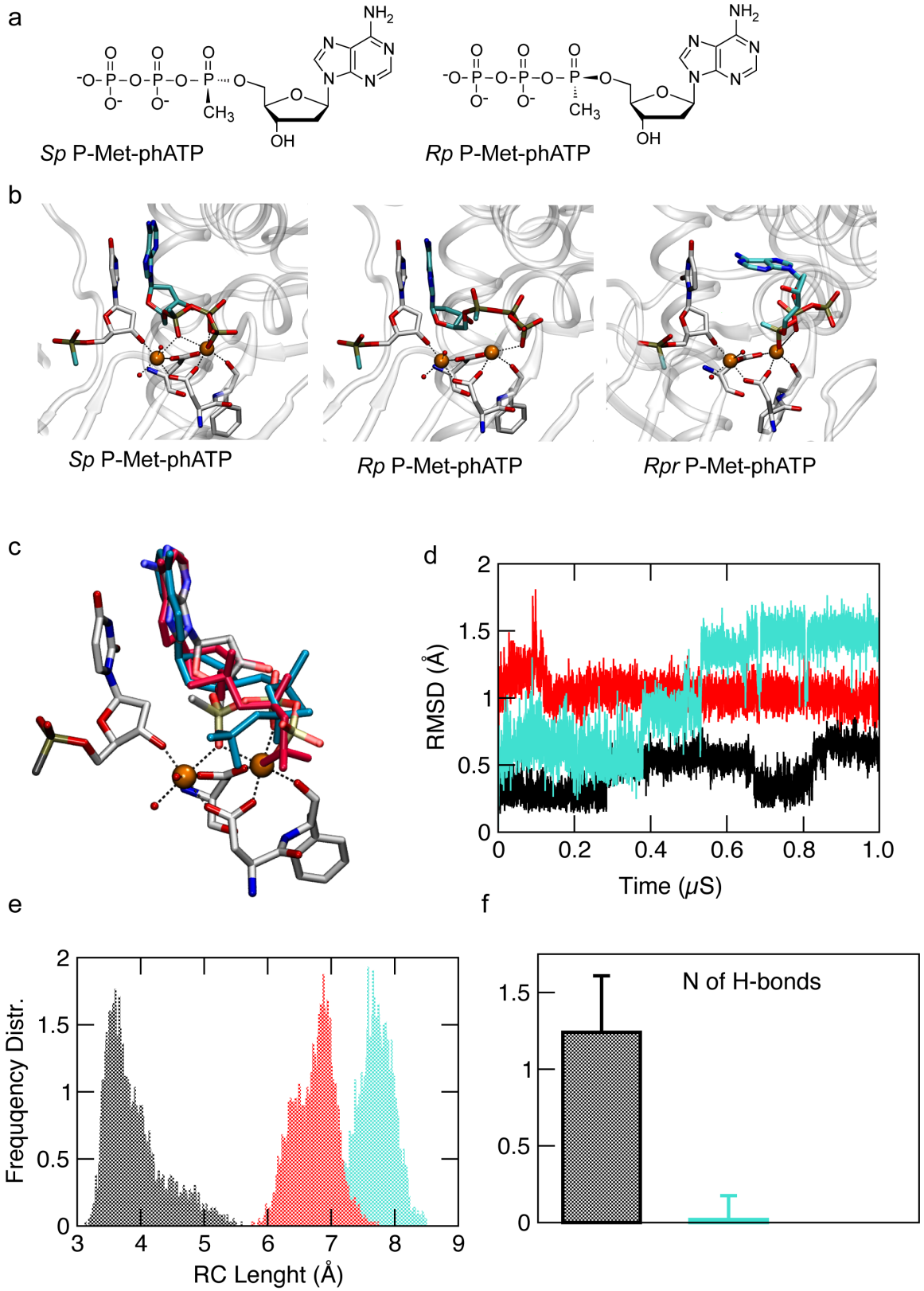
Supplementary Figure 1. Genetic polymer backbones

Examples of different genetic polymer backbones including the canonical polyanionic (red box) phosphodiester backbone of DNA. Variations thereof comprising single atom substitution include the phosphoroamidate DNA (PN-DNA) as well as the phosphorothioate (PS-DNA) and boranophosphate (PB-DNA) DNA in which the substitution is stereogenic (creating a new stereocenter at the P)(dashed box). Also shown are genetic polymers with a polycationic backbone (blue box) comprising Deoxyribonucleic guanidines (DNG) and S-methylthioureas (DNmt) and genetic polymers with uncharged backbones (green box) including the P-alkyl-phosphonates (phNAs) (both uncharged and stereogenic) and Deoxyribonucleic sulfones (DN-SO2).



Supplementary Figure 2. phNTP synthesis and HPLC purification.

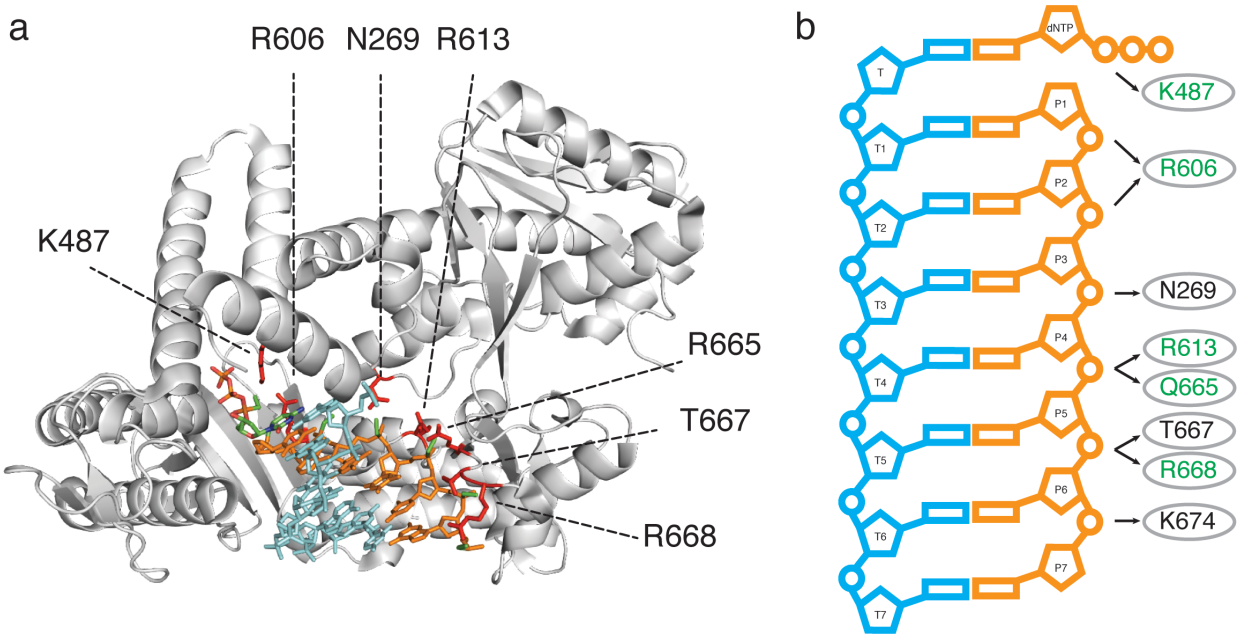
a, *Synthesis scheme for P-α-alkyl-phosphonyl-β,γ-diphosphates (P-α-alkyl-dNTPs, R denotes the alkyl substituent: either Methyl- (CH_3) or Ethyl- (C_2H_5), while the B denotes nucleobases A, T, G, C or U.* ***b,*** *Scheme of the two diastereoisomers (Sp / Rp) of each nucleotide (left panel), HPLC separation of diastereoisomers (middle panel) and incorporation of HPLC peaks 1 and 2 by RT521 polymerase (right panel). Only one of the HPLC peaks is a polymerase substrate.*


a
Sp P-Met-phATP
Rp P-Met-phATP
b
Sp P-Met-phATP
Rp P-Met-phATP
Rpr P-Met-phATP
c
d
RMSD (Å)
Time (μS)
e
Frequqency Distr.
RC Lenght (Å)
f
N of H-bonds

Supplementary Figure 3. Global and structural modelling of Pol / P-Met-phATP diastereoisomer complexes.

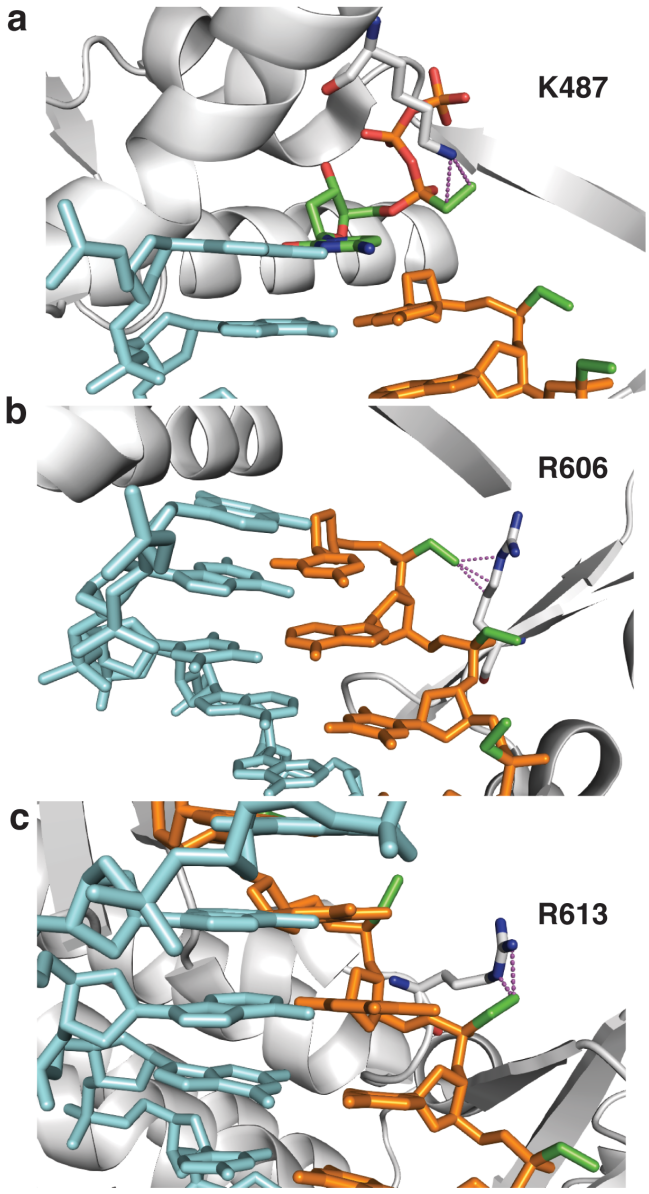
Global and local site structural properties of Polymerase / primer / phATP complex accommodating the incoming different P-Met-phATP diastereoisomers in the polymerase active site starting from the RB69 DNA Polymerase / dNTP / dsDNA ternary holoenzyme complex crystallographically resolved (PDBid 4M3R) at 2.07 Å resolution[7].

***a**, Chemical structure the S_p P-Met-phATP and R_p phATP diastereoisomers. **b**, Snapshot extracted from our MD simulations displaying the incoming S_p P-Met-phATP and R_p / R_pr P-Met-phATP diastereoisomers in the polymerase active site. **c-f**, **Color code**: black identifies Methyl-S_p system, red Methyl-R_p while cyan represents Methyl-R_pr system (see Methods section for nomenclature). **c,** View of the active site of Polymerase / phDNA / P-Met-phATP ternary complex. Carbon atoms are coloured in white, oxygen in red, nitrogen in blue and phosphorus in maroon. Incoming phATP extracted from Met-S_p (coloured accordingly with the atom species), Met-R_p (in red) and Met-R_pr (in cyan) systems are overlapped for comparison. Dashed line indicates Mg-O coordination. **d,** RMSD of the incoming P-Met-phATP (heavy atoms only). **e,** Length of the reaction coordinate describing the nascent phosphodiester bond (i.e. the length between -3'O^- of the 3'-end nucleotide and the αP of the incoming P-Met-phATP. **f,** Population of Watson-Crick H-bond between the incoming P-Met-phATP and its complementary base on the DNA template strand. Results are presented as the mean with standard deviation.*



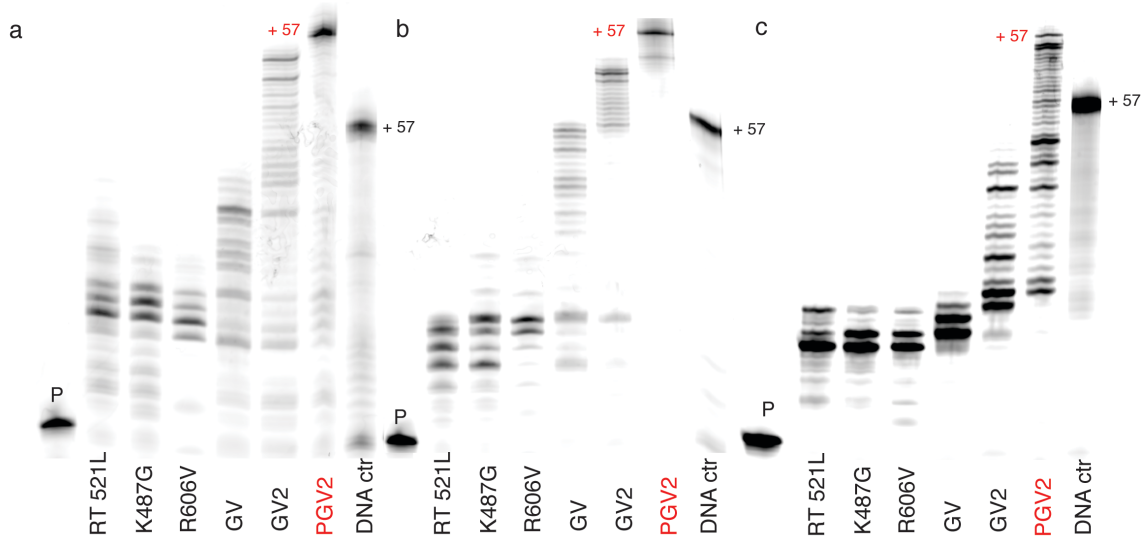
Supplementary Figure 4. Steric clashes with P-ethyl-phNA backbone.

a, *Model of the homologous KOD polymerase (PDB: 4K8Z) bound to a triphosphate substrate and a phNA-DNA primer-template duplex (using substrates from the RB69 structure PDB: 1IG9) the nascent (primer) strand is shown in orange with the ethyl-phosphonate moiety highlighted in green while the template is orange; residues targeted for structure-based saturation mutagenesis are shown in cyan.* ***b,*** *Schematic of the template (cyan) and primer (orange) duplex and the polymerase residues which interact with the phosphate groups on the primer strand or the dNTP.*



Supplementary Figure 5, Steric clashes with the phosphono-ethyl group

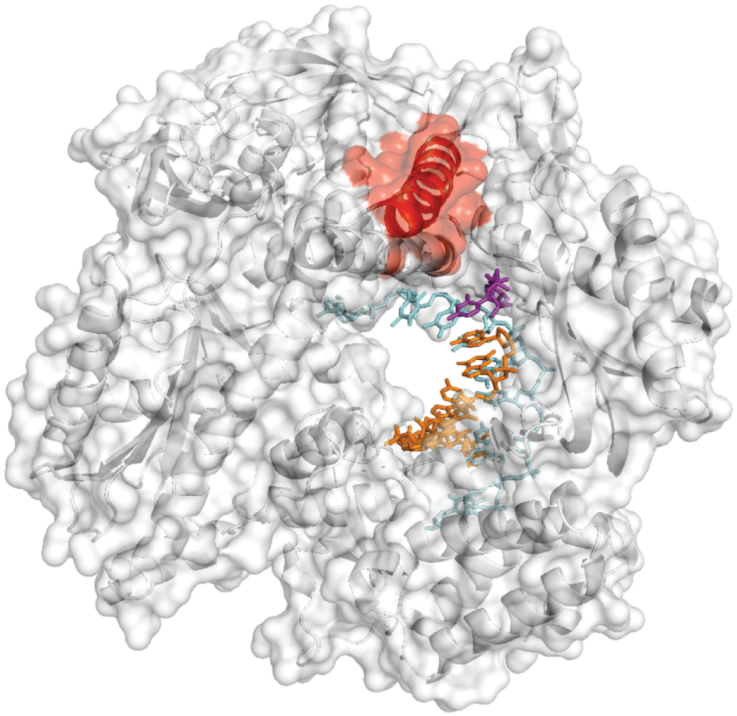
a, *Detailed view of the active site (of the RB69 structure derived primer-template duplex and dNTP overlaid onto the KOD binary structure (as in Supplementary Fig. 4)), showing ethyl group on the triphosphate (in an ideal conformation for catalysis) coming into close contact with K487,* ***b,*** *close contact of R606 and the ethyl moiety in the newly-formed phNA strand,* ***c,*** *close contact of R613 and the +3 residue on the nascent phNA strand. Dashes indicate contacts where the heavy atom distance is disallowed or less than 2.9Å.*



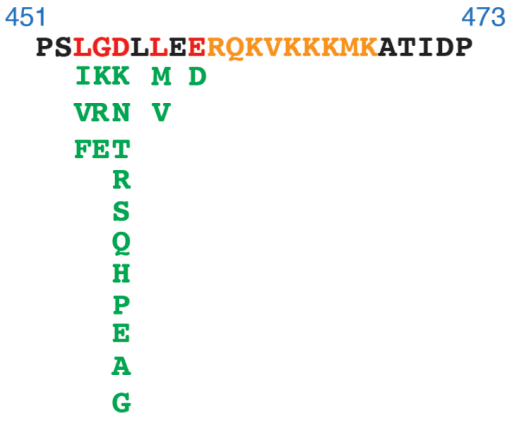
Supplementary Figure 6. P-Et-phNA synthesis on different templates.

a-c, *Primer extension using all P-Et-phNTPs by the parent polymerase RT521L, individual single, double (GV: K487G, R606V) and triple (GV2: K487G, R606V, R613V) mutants and the final quadruple mutant polymerase PGV2 on different DNA templates:* ***a,*** *a template without dA TnT,* ***b,*** *a low purine template THT6 and* ***c,*** *a template with 18/57 purines TBS (Template sequences Supplementary Table 6). While single mutant polymerases show little or no improvement in phNA synthesis, the GV double and GV2 triple mutant show clear improvements, with the PGV2 polymerase showing full length synthesis of 57nt P-Et-phNA oligomers.*

a

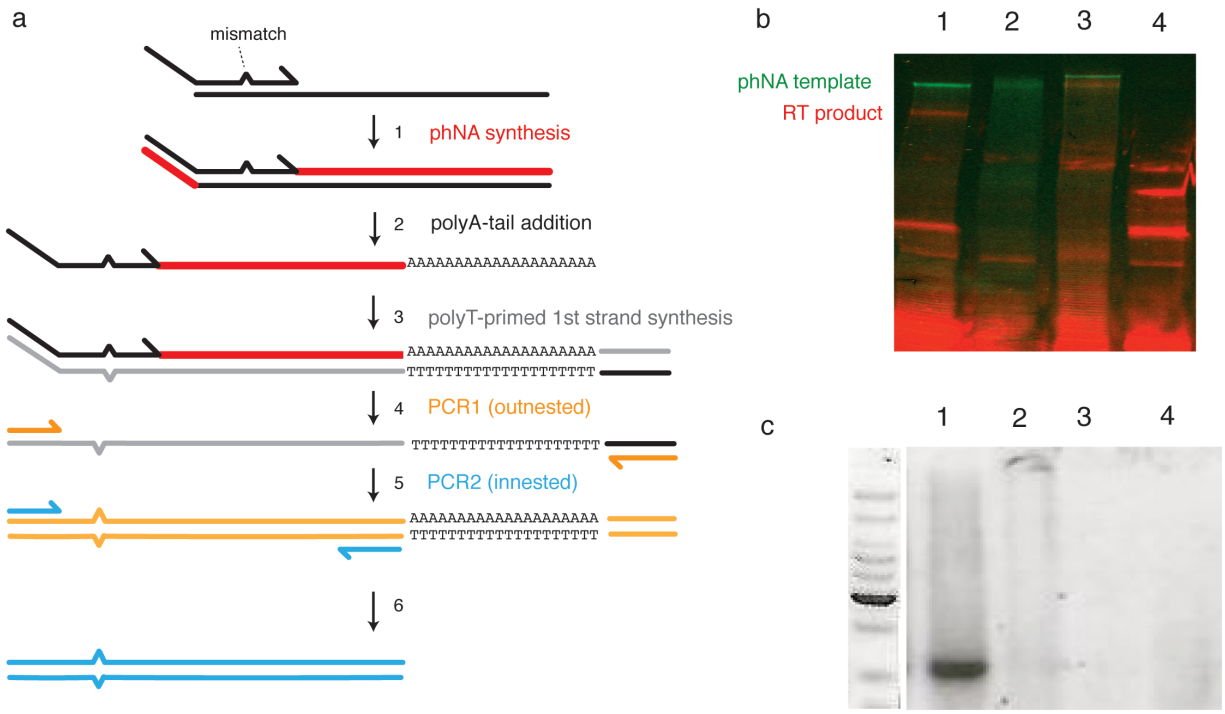


b



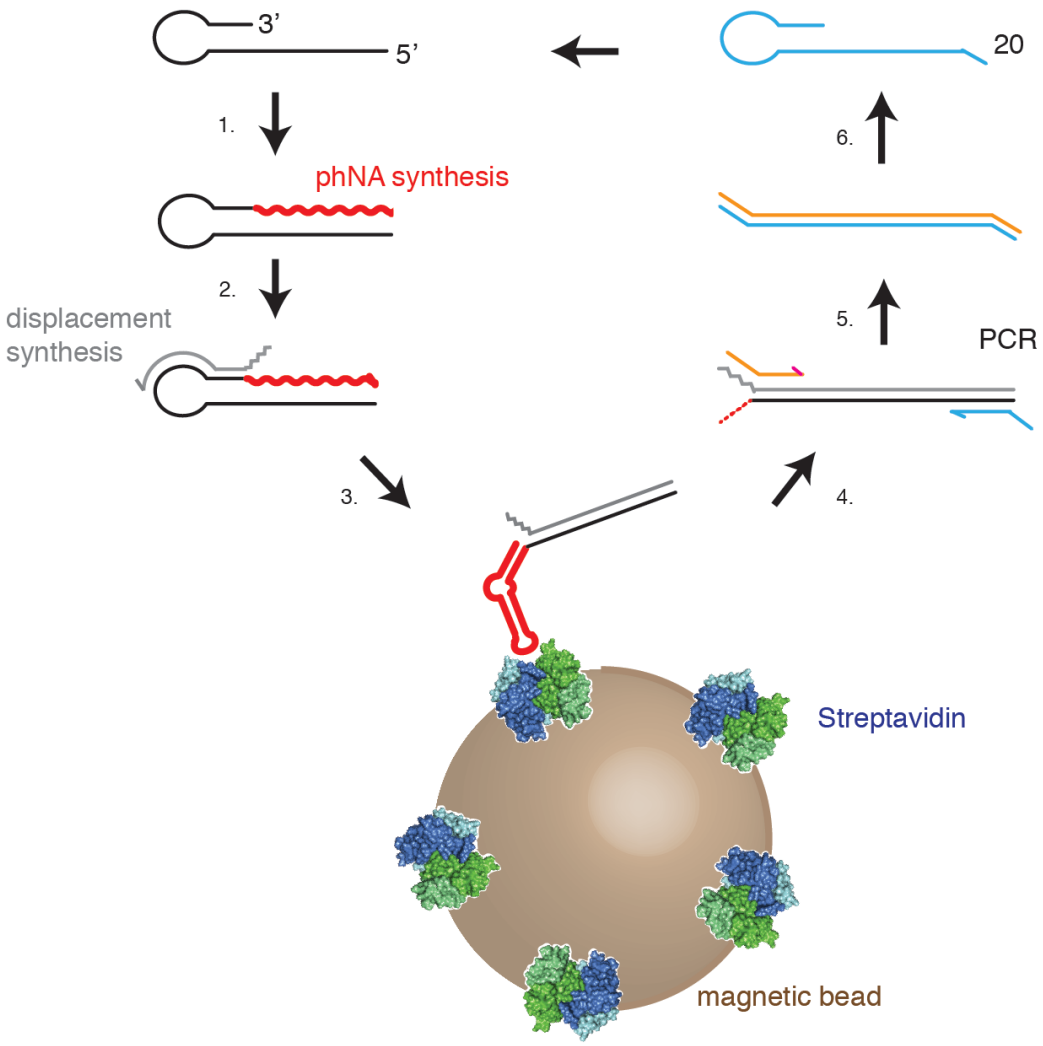
Supplementary Figure 7. CST selection polymerase library.

Motif 5 library is shown. ***a,*** *Polymerase structure with primer (orange), template (cyan), nucleotide triphosphate (violet), and the motif 5 subjected to targeted mutagenesis in red.* ***b,*** *Amino acid sequence of mutated motif 5 residues, black (unchanged), red (randomized) and orange (low mutagenesis spike at 5%). Residues in selected polymerases are shown in green. D455 is the most heavily mutated residue with the D455P mutation found to enhance phNA synthesis in the GV2 context.*



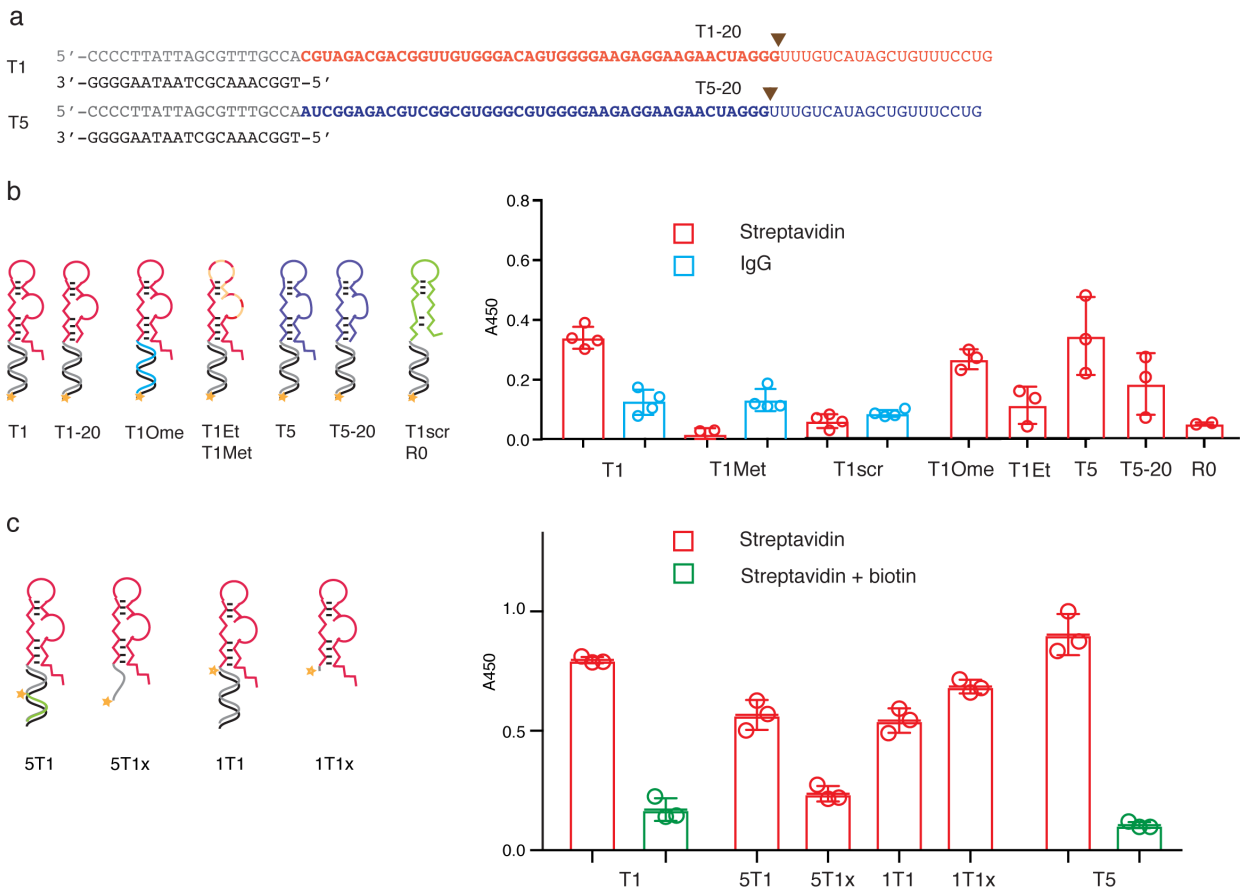
Supplementary Figure 8. Reverse Transcription of phNA into DNA.

a, *Reverse Transcription scheme: 1-2) A DNA primer containing a diagnostic mismatch is used to synthesize phNA against a DNA template, 3-4) After purification of the phNA-DNA chimeric strand, the phNA 3' end is poly-deoxyadenylated and a poly-dT-primer is used to prime 1st strand synthesis against phNA using the RT521L: E664K reverse transcriptase enzyme, 5) purification of the reverse transcription product and double-out-nesting PCR 1, 6) in-nest PCR 2 to prepare sample for cloning, which contains the diagnostic mismatch sequence introduced only if derived from synthesized phNA.* ***b,*** *PAGE of the reverse transcription products from a poly(dT) primer (red) and a phNA template polyadenylated at its 3' end (green): 1) dNTPs, phNA template, 2) no dNTPs, phNA template, 3) just dTTP, phNA template, 4) dNTPs, no phNA template. RT 1st strand synthesis product of correct size from polydT primer is visible only when both phNA template and dNTP substrates are available in the RT reaction (1).* ***c,*** *PCR from the RT reactions described (double outnest). A PCR band of correct size is visible only after successful RT reaction (1).*



Supplementary Figure 9. phNA aptamer selection

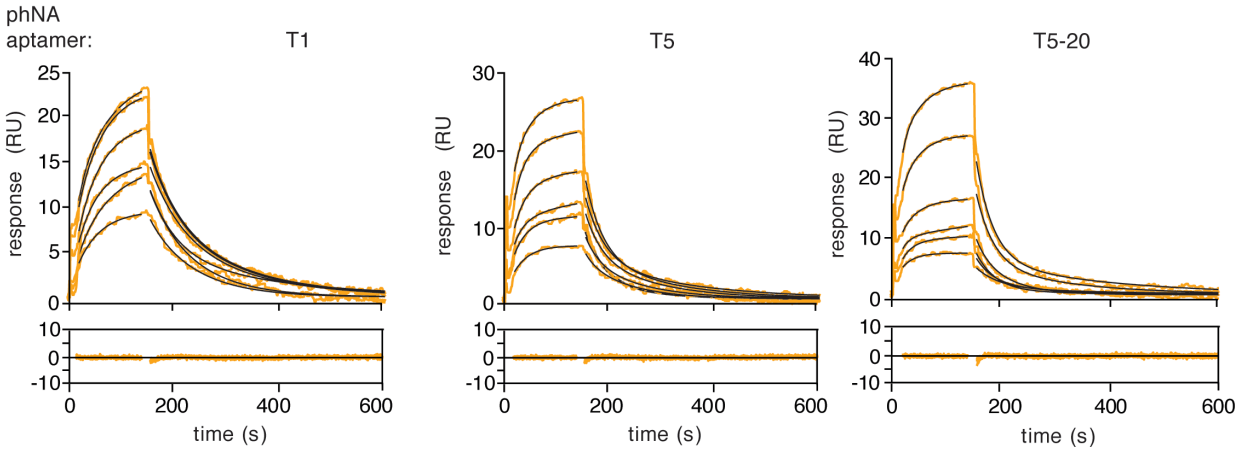
DNA display aptamer selection scheme: 1) a DNA hairpin works as both primer and template for phNA synthesis; 2) the synthesized phNA is then displayed by strand-displacement synthesis with a DNA polymerase filling in the DNA duplex, 3-4) the synthesized and displaced phNA is allowed to fold and bind to streptavidin-coated magnetic beads. Binders are captured and eluted and 5) the encoding DNA is amplified by PCR using specific primers. After processing of the PCR products, 6) the (-) strand is removed by exonuclease digestion, regenerating the DNA hairpin structure ready for the next round of selection.



Supplementary Figure 10. phNA aptamer variants

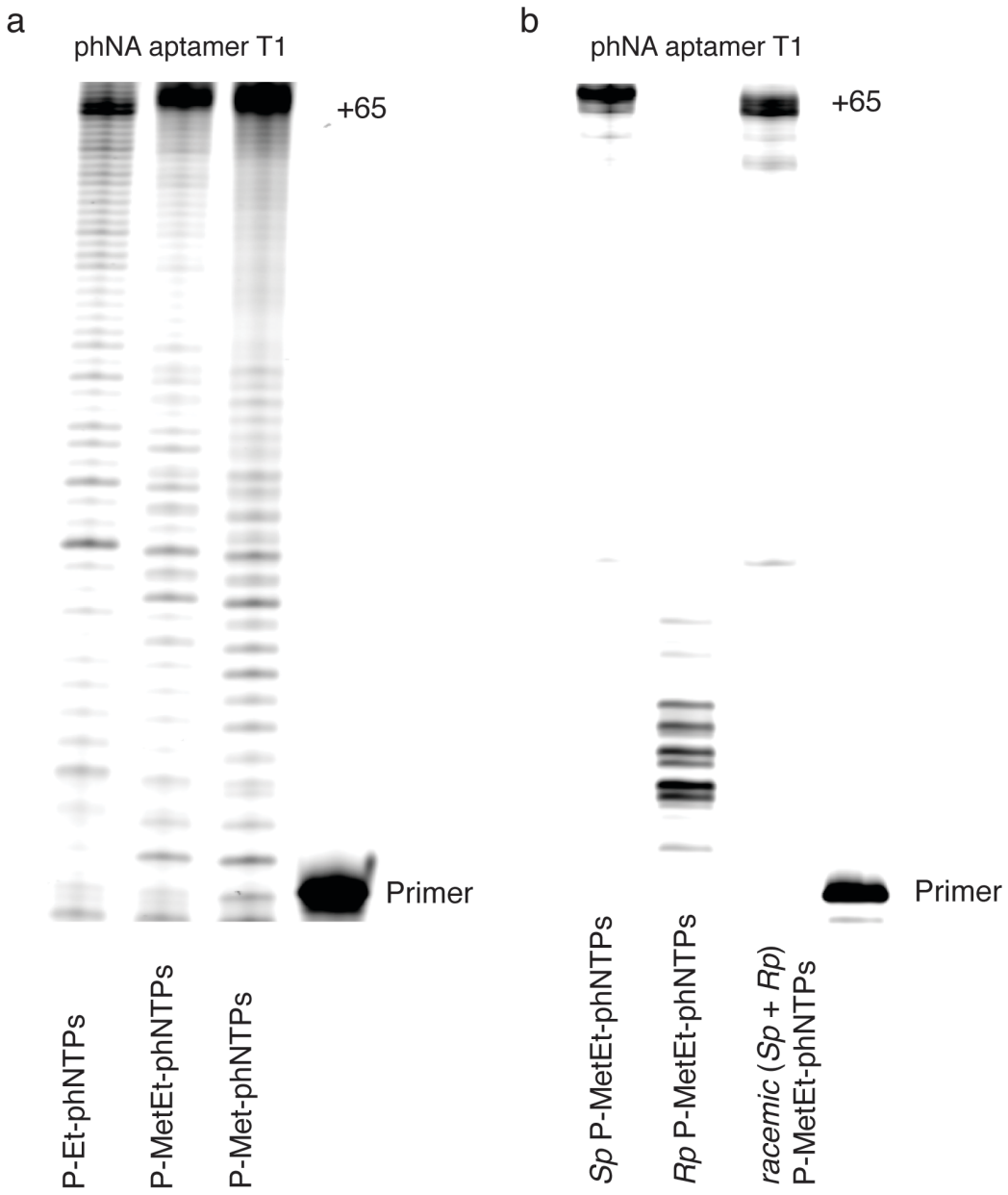
a, *Sequence of aptamer clones T1 and T5, with DNA primer (grey), complementary DNA oligonucleotide (black), phNA portion in red (T1) or blue (T5) with pHNA deriving from random sequence segment (bold) and shortened version T1-20, T5-20 (shortened by 20 nts from 3', black triangle).* ***b,*** *Cartoon representation of different phNA aptamer variants T1, T5, T1-20, T5-20, T1Ome (T1 synthesized from a different sequence 2'OMe primer (cyan)), T1met / T1et (T1 synthesized as all P-Met- resp. P-Et-phNA), T1scr (T1 synthesized with the segment deriving from random sequence (bold) scrambled) and R0 (arbitrary phNA aptamer sequence), all with 5'-FITC-dT residue (yellow star) for detection (left panel). ELONA assay of different phNA aptamer variants (shown in left panel) for binding to Streptavidin (SA) (red) or immunoglobulin (IgG) control (cyan). Results are presented as the mean with standard deviation (n=4 (T1, T1Met, T1scr); n=3)(T1Ome, T1Et, T5, T5-20, R0)).* ***c,*** *Cartoon representation of phNA aptamer variants 5T1 (synthesized from an RNA primer (green) followed by a FITC-dT residue (yellow star) and 4 DNA nucleotides (grey)), 5T1x (5T1 after cleavage of the RNA portion with RNAse, which leaves the FITC-dT residue and 4 DNA nucleotides attached to the T1 phNA aptamer),*

1T1 (synthesized from a DNA primer with a 3'-FITC-dT residue (yellow star)), 1T1x (degradation of 1T1 by BAL-31 nuclease leaves only a single FITC-dTp residue attached to the T1 phNA aptamer). Binding of T1 and T5 phNA aptamers to Streptavidin (red) + / - addition of biotin (green) are shown for reference. Results are presented as the mean with standard deviation (n=3).



Supplementary Figure 11. Characterisation of phNA aptamers

SPR sensograms with residuals of the curves fit of streptavidin binding by the phNA aptamers T1 and T5 as well as the truncated aptamer T5-20. Summary of binding kinetics determined from SPR (Supplementary Table 3).



Supplementary Figure 12. phNA synthesis with different substrates

Synthesis of phNA aptamer T1 (Supplementary Fig. 10a) by polymerase PGV2: ***a,*** *using (left to right) all P-Et-phNTP mix, P-MetEt-phNTP mix or all P-Met-phNTP mix (yielding aptamers T1et, T1 and T1met (Supplementary Fig. 10b)).* ***b,*** *using (left to right) the HPLC-purified mix of peaks that were active polymerase substrates (presumed S_p P-MetEt-phNTP mix)(Supplementary Fig. 2), phNTP mix composed of the non-substrate peaks (presumed R_p P-MetEt-phNTP mix) (Supplementary Fig. 2), and all peaks combined (racemic P-MetEt-phNTP mix).*

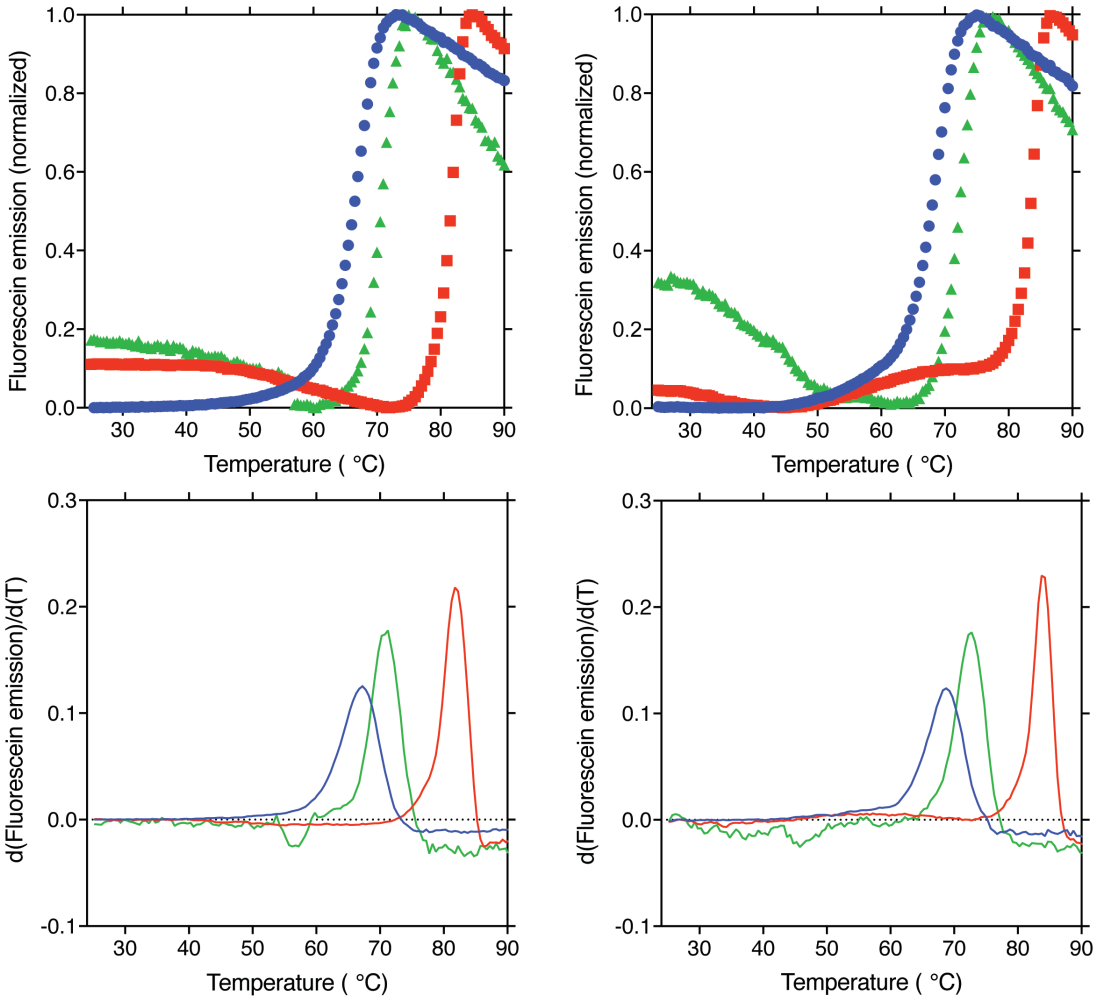
a

DNAΔ
5′-F-CCCCTTATTAGCGTTTGCCA-3′
3′-Q-GGGGAATAATCGCAAACGGTCTACGCTACGTATGCGCTATGTCGTA-5′

DNA-phNA
5′-F-CCCCTTATTAGCGTTTGCCAGATGCGATGCATACGCGATACAGCAT-3′
3′-Q-GGGGAATAATCGCAAACGGTCTACGCTACGTATGCGCTATGTCGTA-5′

DNA-DNA
5′-F-CCCCTTATTAGCGTTTGCCAGATGCGATGCATACGCGATACAGCAT-3′
3′-Q-GGGGAATAATCGCAAACGGTCTACGCTACGTATGCGCTATGTCGTA-5′

b



Supplementary Figure 13. Thermostability

Melting and renaturation profiles of DNA and phNA. ***a,*** *Primer DNA oligonucleotide (blue) on DNA template (black), with synthesized phNA strand (green) and the corresponding synthesized DNA strand (red).* ***b,*** *Hybridization (left) and Denaturation (right) curve showing normalized fluorescence intensity vs Temperature or the respective first derivative of the fluorescence emission vs. temperature (bottom panel). Thermal stability data is summarized in Supplementary Table 4.*

Supplementary Tables

Error Rate per base ($x10^{-4}$)		Substitution			
		A	T	C	G
Expected Base	G	13.11	0.00	0.00	0.00
	A	0.00	0.00	0.00	0.00
	T	8.74	0.00	0.00	4.37
	C	4.37	4.37	0.00	0.00

Supplementary Table 1. Fidelity

Aggregate fidelity of forward synthesis and reverse transcription of phNA.

Sequence	Rank	Abundance (% of total library)
CGTAGAAGACGGTTGTGGGACAGTG	1	84.85
AATCGGAGACGTCGGCGTGGGCGTG	2	1.57
CGACCGGAGCGTTGTGGGGCAGTGG	3	1.33
ATTGAGACAGCGGTGTGGGGCGGTG	4	0.88
ATACGACGCGAGTTGTGGGGCAGTG	5	0.54
TCCTGAGCATGGCACTATGTTTACT	6	0.06
TGAGCCGCTTAAAGCTACCAGTTAT	7	0.05
TCCCCGTACGCCGGGCAATAATGTT	8	0.04
TCGGGTACGCAATCGCCGCCAGTTA	9	0.04
CCATGCCGCTTTTCTTGGCACGATT	10	0.04

Supplementary Table 2. phNA aptamer sequences

Top 10 most abundant sequences from phNA aptamers selection after five rounds of selection are given with abundances (in %), the sequences shown correspond to the randomized sequence segment of the aptamer library. Rank 1 sequence corresponds to aptamer T1, Rank2 sequence corresponds to aptamer T5.

Aptamer	k_{a1} (M^{-1} s^{-1})	k_{d1} (s^{-1})	K_{D1} (M)	k_{a2} (M^{-1} s^{-1})	k_{d2} (s^{-1})	K_{D2} (M)
T1	$2.65x10^{5}$	0.0188	$7.09x10^{-8}$	$9.84x10^{4}$	0.012	$1.21x10^{-7}$
T1-20	$7.83x10^{4}$	0.0075	$9.56x10^{-8}$	$7.99x10^{4}$	0.019	$2.39x10^{-7}$
T5	$7.30x10^{4}$	0.0316	$4.32x10^{-7}$	$1.42x10^{5}$	0.013	$8.78x10^{-8}$
T5-20	$2.02x10^{5}$	0.0104	$5.16x10^{-8}$	$3.57x10^{5}$	0.028	$7.93x10^{-8}$

Supplementary Table 3. Kinetic measurements SPR

Binding kinetics of phNA aptamers binding to streptavidin using SPR. Biphasic kinetics revealing K_D for T1: 71-121nM

T1-20: 96 - 239nM

T5: 88 - 432nM

T5-20: 52 - 79nM)

	Hybridisation		Denaturation	
	T_m (^oC)	T_{max} (^oC)	T_m (^oC)	T_{max} (^oC)
Primer Alone	65.58	67.25	67	68.75
primer-phNA chimera	70.16	71.25	72.05	71.25
DNA-DNA duplex	81.42	81.75	83.03	83.75

Supplementary Table 4. Thermostability

Thermal stability (^oC) as determined for DNA gapped duplex, DNA-phNA x DNA heteroduplex and DNA x DNA homoduplexes as described in Supplementary Fig. 12. Note that as FITC fluorescence decreases with increasing temperature[18], T_{max} is a more useful parameter than T_m in this case. Thermal stabilization (DT compared to primer alone is 3 - 4 ^oC (0.11 - 0.15 ^oC / base) for phNA and is 14.5 - 15.0 ^oC (0.54-0.58 ^oC / base) for DNA.

Pol	dNTP	V_{max}, %/min	K_m, μM	V_{max}/K_m, % μM^{-1} min^{-1}	k_{cat}, min^{-1}	k_{cat}/K_m, μM^{-1} min^{-1}
RT521L	dA	9.5 ± 0.6	0.40 ± 0.1	48 ± 9.6	47 ± 3	120
	phA	4.9 ± 0.3	39 ± 9.6	0.33 ± 0.06	24 ± 1.5	0.63
PGV2	dA	11.4 ± 0.7	1.38 ± 0.17	9.3 ± 0.7	63 ± 4	46
	phA	11.6 ± 0.3	82 ± 9.3	0.15 ± 0.01	64 ± 2	0.79
RT521L	dT	10.8 ± 0.6	1.6 ± 0.8	15.8 ± 3.0	54 ± 3	33
	phU	7.5 ± 1.9	53 ± 16	0.24 ± 0.06	37 ± 9	0.71
PGV2	dT	11.5 ± 1.9	5.5 ± 2.2	4.9 ± 1.3	61 ± 7	11.9
	phU	7.7 ± 0.8	239 ± 46	0.07 ± 0.011	42 ± 4	0.18
RT521L	dC	11.7 ± 0.5	0.19 ± 0.04	93 ± 12	59 ± 2	308
	phC	5.8 ± 0.4	7.6 ± 0.4	0.76 ± 0.01	29 ± 2	3.83
PGV2	dC	13.2 ± 1.8	1.8 ± 0.4	10.3 ± 1.2	73 ± 9	41
	phC	18.5 ± 1.6	76 ± 13	0.25 ± 0.01	103 ± 8	1.35
RT521L	dG	4.9 ± 0.37	0.18 ± 0.05	31 ± 7.9	25 ± 2	135
	phG	4.4 ± 0.6	5.4 ± 1.4	0.93 ± 0.20	22 ± 3	4.11
PGV2	dG	4.5 ± 0.7	0.87 ± 0.24	5.98 ± 1.13	25 ± 4	29
	phG	8.5 ± 1.0	64 ± 6.9	0.13 ± 0.004	47 ± 5	0.73

Supplementary Table 5. Steady-state kinetics of single nucleotide incorporation.

Values for V_{max}, K_m, k_{cat} for polymerases RT521L and PGV2 and all four dNTP and P-MetEt-phNTPs (phA: P-Et-phATP, phU: P-Met-phUTP, phC: P-Met-phCTP, phG: P-Et-phGTP) are given.

Oligonucleotide Nam	Sequence
TempN	CTCACGATGCTGGACCAGATAAGCACTTAGCCACGTAGTGCTGTTCGGTAATCGATC TGGCAAACGCTAATAAGGGG
TnT	CTCCGTGCTGGCCGTGCCTTGCCCGTGTGCTGTTCGGTTCGTCTGGCAAACGCTAA TAAGGGG
TnC	CTCACATCTACCAATAACACTTACCACTATCTTTCTAATCATCTGGCAAACGCTAATAA GGGG
THT6	TTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGT TGGCAAACGCTAATAAGGGG
TBW	CTCACTATGCTGCACCTGTTCACCACTTATCCACGTACTGCTGTTCTGTATTCCATCT GGCAAACGCTAATAAGGGG
TBS	TCACTACTGCTACCGTTATCGCCACTATCGCTGTTACTGCTGTTCGCTATCGCTATCT GGCAAACGCTAATAAGGGG
TempT	AAAAAAAAAAAAAAAAATGGCAAACGCTAATAAGGGG
TnD	CCCTCCTTCTTCCTCTTCCCCATGCTGTACCACATACGCACTTATCCACGTCGTGCTG TTCGGTAATCGATCTGGCAAACGCTAATCAGGCTATTATTAATTT
TSCD	ATGCTGTATCGCGTATGCATCGCATCTGGCAAACGCTAATAAGGGG
TNT3	GCCCGCCCCGCGCCCGCGCCCGCCGCCCGCCCTCCGTGCTGGCCGTGCCTTGCC CGTGTGCTGTTCGGCGTCTTTGGCAAACGCTAATGAGGGG
TempN25	ACGTAGTGCTGTTCGGTAATCGATCTGGCAAACGCTAATAAGGGG
L3	CAGGAAACAGCTATGACAAACCC
L3polyT	CAGGAAACAGCTATGACAAATTTTTTTTTTTTTTTTTTTTT
NAPfd	CAGTATCGACAAAGGACCCCTTATTAGCGTTTGCCA

Oligonucleotide Nam	Sequence
FD	CCCCTTATTAGCGTTTGCCA
NAP	CAGTATCGACAAAGGA

Supplementary Table 6.

Oligonucleotides used in primer extension and reverse transcription experiments.

Oligonucleotide Name	Sequence
Hairpin N25	CCCTAGTTCTTCCTCTTCCCNNNNNNNNNNNNNNNNNNNNNNNNTGGCAAACGC TAATAAGGGGAACATCTGCGAGCACGCCATCAGGCCCATCAACCCCCTTATTAGC GTTTGCCA
Test7	CCCTAGTTCTTCCTCTTCCC
AntiHairPin35	AGCCGAGGTCGACAACTGGCAAACGCTAATAAGGGGGTTGATGGGC CTGAT
HPOut	AGCCGAGGTCGACAACTGG
BsaIOutnest	GAGTAGGTCTCAACTATGGCAAACGCTAATAAGGGG
DeepSeqRev	CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTC CGATCTCCCCTTATTAGCGTTTGCCA
DeepSeq R0	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGAT CTNNNGGCTACCCCTCTTTCTTCCTCTTCCC
DeepSeq R3	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGAT CTNNNGATCAGCCCTCTTTCTTCCTCTTCCC
DeepSeq R4	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGAT CTNNNACAGTGCCCTCTTTCTTCCTCTTCCC
DeepSeq R5	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGAT CTNNNACAGTGCCCTCTTTCTTCCTCTTCCC
T1 template	CAGGAAACAGCTATGACAAACCCTAGTTCTTCCTCTTCCCCACTGTCCCACAACCG TCGTCTACGTGGCAAACGCTAATAAGGGG
T1-20 template	CCCTAGTTCTTCCTCTTCCCCACTGTCCCACAACCGTCGTCTACGTGGCAAACGCT AATAAGGGG

T5 template	CAGGAAACAGCTATGACAAACCCTAGTTCTTCCTCTTCCCCACGCCCACGCCGAC GTCTCCGATTGGCAAACGCTAATAAGGGG
T5-20 template	CCCTAGTTCTTCCTCTTCCCCACGCCCACGCCGACGTCTCCGATTGGCAAACGCT AATAAGGGG
R0 Template	CAGGAAACAGCTATGACAAACCCTAGTTCTTCCTCTTCCCACATCCGCGTGTCAGC CCAAACTATTGGCAAACGCTAATAAGGGG
SFP (2'O-methyl)	[5'FAM]AGCAGCACAGAGGTCAGATG
SFP minus	CATCTGACCTCTGTGCTGCT
P2dT FITC	AACCTTCGGTAATCGATCTGG[dT-FITC]
FITC FD(-)	TGGCAAACGCTAATAAGGGG[3'FAM]

Supplementary Table 7.

Oligonucleotides used in aptamer selection and characterization experiments

Oligonucleotide Name	Sequence
K487X Fo	GAGTAGGTCTCCTGATCNNSATCCTTGCTAATAGCTTC
K487X Ba	GAGTAGGTCTCGATCAGTCGTTGCCTGTA
R606X Fo	GAGTAGGTCTCACGACGNNSGGGCTTGAAATAGTT
R606X Ba	GAGTAGGTCTCCGTCGTTATCTTGTCCTCCTC
R613X Fo	GAGTAGGTCTCGTTAG,GNNSGACTGGAGCGAG
R613X Ba	GAGTAGGTCTCCCTAA,CTATTTCAAGCCCGCG
T667X Fo	GAGTAGGTCTCCAGATANNSCGCGACCTGAAGGACTAC
Q665X Fo	GAGTAGGTCTC TATCTGCTCGTAGATGAC
Q665X Ba	GAGTAGGTCTC CTCGTAGATGACCAGCTT
N269X Fo	GAGTAGGTCTCACGATTNNSCTCCCCACTTACACCCTT
N269X Ba	GAGTAGGTCTCAATCGTTCTCCTAATGAC
D455P Fo	GAGTAGGTCTCCTCGGTGACCTCTTGGAGGAGAGACAG
D455P Ba	GAGTAGGTCTCACCGAGGAGGCTTGGGAT
R613V Fo	GACTCAGGTCTCGTTAGGGTTGACTGGAGCGAGATA
R613V Ba	GACTCAGGTCTCCCTAACTATTTCAAGCCCGCGCGT
Motif5ba	GAGTCAGGTCTCGAGGASAGACAGAAGGTAAAGAAGAAGATGAAGGCCACGATCG ACCCA
Motif5fo	GAGTCAGGTCTCCTCCTCCANGAGWBBTYYGANGAGGCT

578fo	GAGTCAGGTCTCGCTTCCTCAAGGTCGTCAAG
582ba	GAGTCAGGTCTCGGAAGCGCTTTATCATCTCCTTCTCG
308fo	GAGTCAGGTCTCGCCATTTTTCACTTCACAG
304ba	GAGTCAGGTCTATGGCGCACATTGTGCGACATTTTTTTTGTCTGCC

Supplementary Table 8.

Oligonucleotides used in polymerase engineering.

References

1. Arzumanov AA, Dyatkina NB. An Alternative Route for Preparation of α-Methylphosphonyl-β, γ-Diphosphates of Thymidine Derivatives. *Nucleosides and Nucleotides* 1994, **13**(5)**:** 1031-1037.

2. Pinheiro VB, Taylor AI, Cozens C, Abramov M, Renders M, Zhang S*, et al.* Synthetic genetic polymers capable of heredity and evolution. *Science* 2012, **336**(6079)**:** 341-344.

3. Cozens C, Pinheiro VB, Vaisman A, Woodgate R, Holliger P. A short adaptive path from DNA to RNA polymerases. *Proc Natl Acad Sci USA* 2012, **109**(21)**:** 8067-8072.

4. Sano T, Cantor CR. Cooperative biotin binding by streptavidin. Electrophoretic behavior and subunit association of streptavidin in the presence of 6 M urea. *J Biol Chem* 1990, **265**(6)**:** 3369-3373.

5. Creighton S, Bloom LB, Goodman MF. Gel fidelity assay measuring nucleotide misinsertion, exonucleolytic proofreading, and lesion bypass efficiencies. *Method Enzymol* 1995, **262:** 232-256.

6. Boosalis MS, Petruska J, Goodman MF. DNA-Polymerase Insertion Fidelity - Gel Assay for Site-Specific Kinetics. *J Biol Chem* 1987, **262**(30)**:** 14689-14696.

7. Xia S, Konigsberg WH. Mispairs with Watson-Crick base-pair geometry observed in ternary complexes of an RB69 DNA polymerase variant. *Protein Sci* 2014, **23**(4)**:** 508-513.

8. Wang JM, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general amber force field. *J Comput Chem* 2004, **25**(9)**:** 1157-1174.

9. Singh UC, Kollman PA. An Approach to Computing Electrostatic Charges for Molecules. *J Comput Chem* 1984, **5**(2)**:** 129-145.

10. Berendsen HJC, Vanderspoel D, Vandrunen R. Gromacs - a Message-Passing Parallel Molecular-Dynamics Implementation. *Comput Phys Commun* 1995, **91**(1-3)**:** 43-56.

11. Grest GS, Kremer K. Molecular-Dynamics Simulation for Polymers in the Presence of a Heat Bath. *Phys Rev A* 1986, **33**(5)**:** 3628-3631.

12. Dal Peraro MD, Spiegel K, Lamoureux G, De Vivo M, DeGrado WF, Klein ML. Modeling the charge distribution at metal sites in proteins for molecular dynamics simulations. *J Struct Biol* 2007, **157**(3)**:** 444-453.

13. Genna V, Colombo M, De Vivo M, Marcia M. Second-Shell Basic Residues Expand the Two-Metal-Ion Architecture of DNA and RNA Processing Enzymes. *Structure* 2018, **26**(1)**:** 40-50.

14. Genna V, Gaspari R, Dal Peraro M, De Vivo M. Cooperative motion of a key positively charged residue and metal ions for DNA replication catalyzed by human DNA Polymerase-eta. *Nucleic Acids Res* 2016, **44**(6)**:** 2827-2836.

15. Dreyer J, Brancato G, Ippoliti E, Genna V, De Vivo M, Carloni P*, et al.* First Principles Methods in Biology: From Continuum Models to Hybrid Ab initio Quantum Mechanics/Molecular Mechanics. *Rsc Theor Comput Che* 2017, **9:** 294-339.

16. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of Simple Potential Functions for Simulating Liquid Water. *J Chem Phys* 1983, **79**(2)**:** 926-935.

17. Humphrey W, Dalke A, Schulten K. VMD: Visual molecular dynamics. *J Mol Graph Model* 1996, **14**(1)**:** 33-38.

18. You Y, Tataurov AV, Owczarzy R. Measuring Thermodynamic Details of DNA Hybridization Using Fluorescence. *Biopolymers* 2011, **95**(7)**:** 472-486.