Electronic Supplementary Material (ESI) for ChemComm.

Artificial nucleotide codons for enzymatic DNA synthesis

Nazarii Sabat,[a] Andreas Stämpfli,[b] Marie Flamme,[a] Steven Hanlon,[c] Serena Bisagni,[c] Filippo Sladojevich,[b] Kurt Püntener[c] and Marcel Hollenstein[a]*

[a] Institut Pasteur, Université Paris Cité, CNRS UMR3523, Department of Structural Biology and Chemistry, Laboratory for Bioorganic Chemistry of Nucleic Acids, 28, rue du Docteur Roux, 75724 Paris Cedex 15, France

[b] Pharma Research and Early Development, Roche Innovation Center Basel, F. Hoffmann-La Roche Ltd, Grenzacherstrasse 124, 4070, Basel, Switzerland

[c] Pharmaceutical Division, Synthetic Molecules Technical Development, Process Chemistry & Catalysis, F. Hoffmann-La Roche Ltd, 4070 Basel, Switzerland

***** Corresponding Author: marcel.hollenstein@pasteur.fr

Contents

1. Materials and methods

All reactions were performed under argon in flame-dried glassware. Anhydrous solvents and reagents for reactions were purchased from Sigma Aldrich and Alfa Aesar. Thymidine on solid support (5'-*O*-DMT-thymidine-3'-lcaa-CPG) was ordered from Biosynth. Thymidine phosphoramidite was purchased from ChemGenes and LNA-T phosphoramidite was obtained from Roche. NMR spectra were recorded on a Bruker Avance 500 spectrometer (500.1 MHz for ^{1}H, 125.8 MHz for ^{13}C, and 202.5 MHz for ^{31}P), 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. The NMR signals were assigned using a combination of ^{1}H and ^{31}P 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/H_2O 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 preparative ion exchange column (13 µm, 250 x 22.0 mm). Oligonucleotides were purchased from Microsynth or Integrated DNA Technologies (IDT) companies. All the DNA and RNA polymerases (TdT, PUP, Hemo KlenTaq, *Taq*, Bst 2.0, *Sulfolobus* DNA Polymerase IV (Dpo4), Vent (*exo*$^{-}$), Deep Vent (*exo*$^{-}$), Klenow (*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 Amersham Typhoon phosphorimager with the ImageQuantTL software (v.10.2) (both from Cytiva). The Amicon centrifuge filters (10 kDa) were purchased from Sigma Aldrich. Chromatographic separations for LCMS experiments were performed on a Thermo Scientific™ Vanquish™ Flex Binary UHPLC system (Thermo Fisher Scientific, Reinach, Switzerland). The column used for all separations was a Waters Aquity Premier BEH C18 Peptide 2.1*50mm 1.7µm 300A (Waters (CH) AG, Baden-Dättwil, Switzerland).

2. Experimental section.

2.1 Synthesis of trinucleotide triphosphates. General procedure.

5′-DMTr-protected thymidine on solid support **1** (1067.5 mg, 0.05 mmol) was loaded on the short column attached to the collecting waste flask under argon. All reagent solutions were flushed through the column by using controlled vacuum under argon flow (see Figure SI1). The synthesis started with the preparation of trinucleotide 5′-d(TTT)-3′ **3a**, followed by triphosphorylation, and finished with simultaneous global deprotection and solid support removal (see Figure 2).

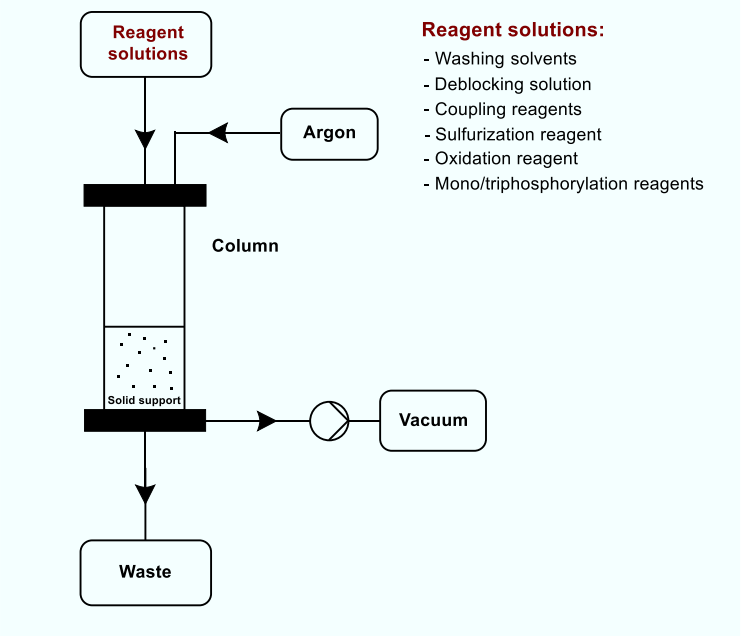


Figure SI1. Schematic representation of manual column synthesis device.

Prior to the synthetic sequence, all of the following reagent solutions, as well as solvents, were prepared and kept dried under argon and molecular sieves:

Deblocking solution (1X): TFA 3% (0.3 mL TFA in 9.7 ml MeCN);

Coupling solution (1X): Phosphoramidite reagent 100 mM (0.5 mmol (10 equiv.) in 5 mL MeCN);

Coupling activator solution (1X): 5-(Ethylthio)-*1H*-tetrazole (ETT), 250 mM (0.5 mmol (10 equiv.) in 2 mL MeCN);

Oxidation solution (1X): I_2 (10 equiv.) 127 mg in 5 mL THF/Py/H_2O (7:2:1);

Sulfurization solution (1X): Beaucage reagent (3*H*-1,2-Benzodithiol-3-one 1,1-dioxide) (10 equiv.) 100 mg in 5 mL of Py;

Monophosphorylation solution: 2-Chloro-1,3,2-benzodioxaphosphorin-4-one (10 equiv.) 101 mg in 5 mL MeCN;

Triphosphorylation solution: Tributylammonium pyrophosphate (10 equiv.) 274 mg in 3 mL of Bu_3N/DMF (1:2).

The synthetic sequence (Figure 2) was performed as follows:

1. Thymidine on solid support **1** (0.05 mmol) was washed with 10-20 mL of MeCN.
2. *1st Deblocking* step to remove DMTr group. The column was slowly (10 min) flushed with 3% TFA solution in MeCN (10 mL). Then carefully washed with 10-20 mL of MeCN.
3. *1st Coupling* step. Nucleoside phosphoramidite reagent (0.5 mmol in 5 mL of MeCN, 100 mM solution) was mixed with ETT activator (0.5 mmol in 2 mL of MeCN, 250 mM solution) under argon and subsequently added to the column. The column was slowly (10 min) flushed with this solution. Then washed with 10-20 mL of MeCN.
4. *Oxidation or Sulfurization*. The column was flushed with the *oxidation solution* (5 mL of I_2 solution in THF/Py/H_2O (7:2:1) or *sulfurization solution* (5 mL of Beaucage reagent solution in Py). Then washed with 10-20 mL of MeCN.
5. *2nd Deblocking* step to remove DMTr group. The column was slowly (10 min) flushed with 3% TFA solution in MeCN (10 mL). Then carefully washed with 10-20 mL of MeCN.
6. *2nd Coupling* step. Nucleoside phosphoramidite reagent (0.5 mmol in 5 mL of MeCN, 100 mM solution) was mixed with ETT activator (0.5 mmol in 2 mL of MeCN, 250 mM solution) under argon and subsequently added to the column. The column was slowly (10 min) flushed over with this solution. Then washed with 10-20 mL of MeCN.
7. *Oxidation or Sulfurization*. The column was flushed with the *oxidation solution* (5 mL of I_2 solution in THF/Py/H_2O (7:2:1) or *sulfurization solution* (5 mL of Beaucage reagent solution in Py). Then washed with 10-20 mL of MeCN.
8. *3rd Deblocking* step to remove DMTr group. The column was slowly (10 min) flushed with 3% TFA solution in MeCN (10 mL). Then carefully washed with 10-20 mL of MeCN.
   Monophosporylation step. The column was slowly (10 min) flushed with 5 mL of Py/MeCN (1:5). Then, slowly (10 min) flushed with *monophosphorylation solution* (2-Chloro-1,3,2-benzodioxaphosphorin-4-one in 5 mL MeCN);
   Triphosphorylation step. The column was washed with 5 mL of DMF. Then slowly (10 min) flushed with *triphosphorylation solution* (Tributylammonium pyrophosphate in 3 mL of Bu_3N/DMF (1:2)). Then washed with 10-20 mL of MeCN.
9. *Oxidation.* The column was flushed with the *oxidation solution* (5 mL of I_2 solution in THF/Py/H_2O (7:2:1). Then washed with 10-20 mL of MeCN.
10. *Global deprotection and solid support removal*. The trinucleotide triphosphate on solid support was removed from the column and stirred for 2 hours in the mixture of aqueous NH_4OH/$MeNH_2$ (10 mL, 1:1). The mixture was filtered on sintered glass filters and re-filtered *via* a syringe filter, then concentrated under vacuum.
11. *HPLC Purification*. The crude was re-dissolved in water (2 mL) and purified by HPLC using a preparative ion exchange column (Buffer A: 10 mM TEAB to Buffer B: 1 M TEAB). Finally, the collected aqueous solution was freeze-dried, yielding the corresponding trinucleotide triphosphate as yellowish solids.

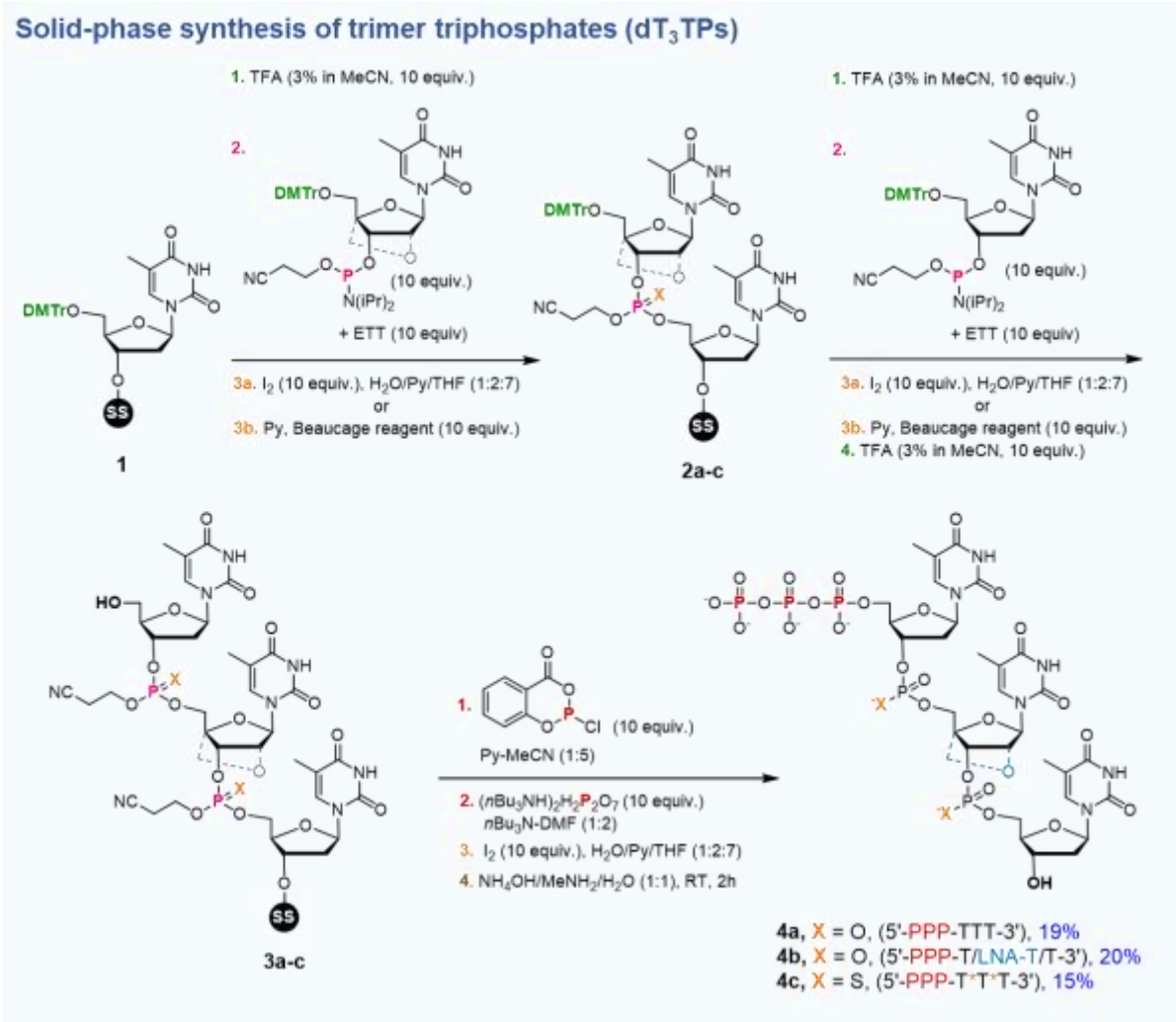
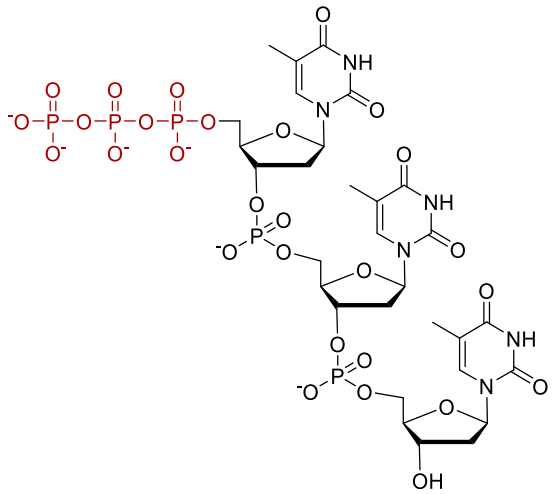


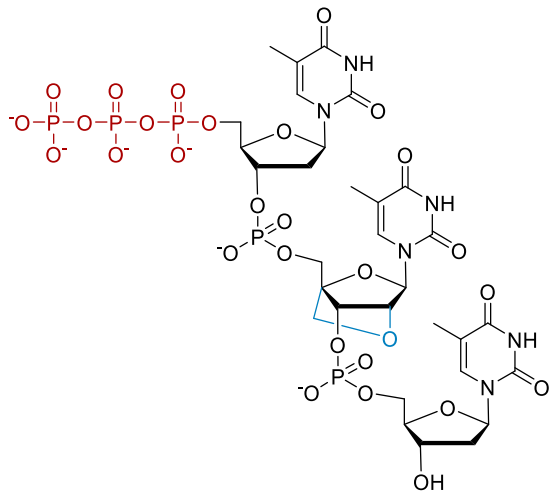
Figure 2. Solid-phase synthesis of trimer triphosphates (dT_3TPs).

Trinucleotide-5′-*O*-triphosphate 4a



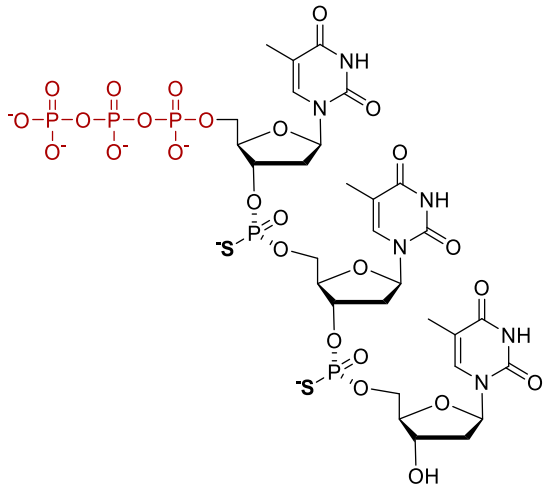
Trinucleotide triphosphate **4a** (10.4 mg, 10 µmol, 19%) was prepared according to the general procedure. ^{1}H NMR (500 MHz, D_2O) 7.57 (d, *J* = 1.0 Hz, 1H), 7.55 (d, *J* = 1.5 Hz, 1H), 7.51 (d, *J* = 1.0 Hz, 1H), 6.19 – 6.12 (m, 3H), 4.27 – 4.24 (m, 2H), 4.19 – 4.16 (m, 2H), 4.07 – 4.05 (m, 2H), 4.02 – 3.94 (m, 6H), 2.38 – 2.36 (m, 3H), 2.23 – 2.20 (m, 3H), 1.78 (d, *J* = 1.2 Hz, 3H), 1.76 (d, *J* = 1.1 Hz, 3H), 1.74 (d, *J* = 1.1 Hz, 3H). ^{31}P NMR (203 MHz, D_2O) δ -1.23 (s,1P); -1.36 (s, 1P), -6.28 (d, *J* = 18.5 Hz, 1P), -11.57 (d, *J* = 19.1 Hz, 1P), -22.19 (m, 1P). HRMS calcd. for $C_{30}H_{41}N_6O_{28}P_5$ $[M-2H]^{2-}$: 1088.0668, found: 1088.0680.

LNA-trinucleotide-5′-*O*-triphosphate 4b



Trinucleotide triphosphate **4b** (11.2 mg, 10 µmol, 20%) was prepared according to the general procedure. ^{1}H NMR (500 MHz, D_2O) 7.65 (d, *J* = 1.3 Hz, 1H), 7.63 (d, *J* = 1.5 Hz, 1H), 7.50 (d, *J* = 1.7 Hz, 1H), 6.19 – 6.16 (m, 2H), 5.53 (s, 1H), 4.84 – 4.82 (m, 1H), 4.56 (s, 1H), 4.42 – 4.39 (m, 2H), 4.32 (d, *J* = 5.5 Hz, 1H), 4.30 – 4.29 (m, 1H), 4.26 – 4.23 (m, 2H), 4.08 – 4.06 (m, 2H), 4.04 – 4.01 (m, 2H), 3.98 – 3.93 (m, 2H), 2.42 – 2.37 (m, 3H), 2.24– 2.19 (m, 1H), 2.15– 2.10 (m, 1H), 1.77 (d, *J* = 1.0 Hz, 3H), 1.73 (d, *J* = 1.2 Hz, 3H), 1.69 (d, *J* = 1.1 Hz, 3H). ^{31}P NMR (203 MHz, D_2O) δ -1.32 (s,1P); -2.00 (s, 1P), -9.73 (m, 1P), -11.63 (d, *J* = 19.2 Hz, 1P), -22.68 (m, 1P). HRMS calcd. for $C_{31}H_{41}N_6O_{29}P_5$ $[M-2H]^{2-}$: 1116.0617, found: 1116.0622.

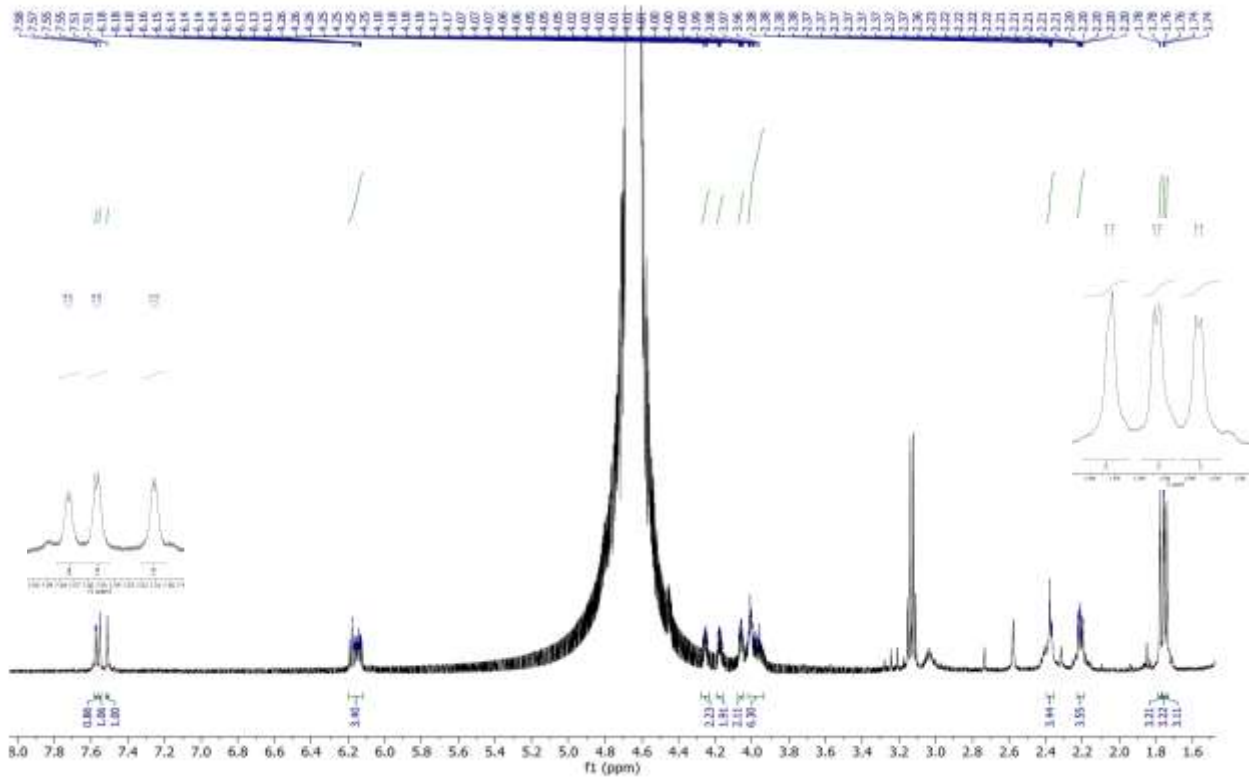
Phosphorothioate-trinucleotide-5′-*O*-triphosphate 4c



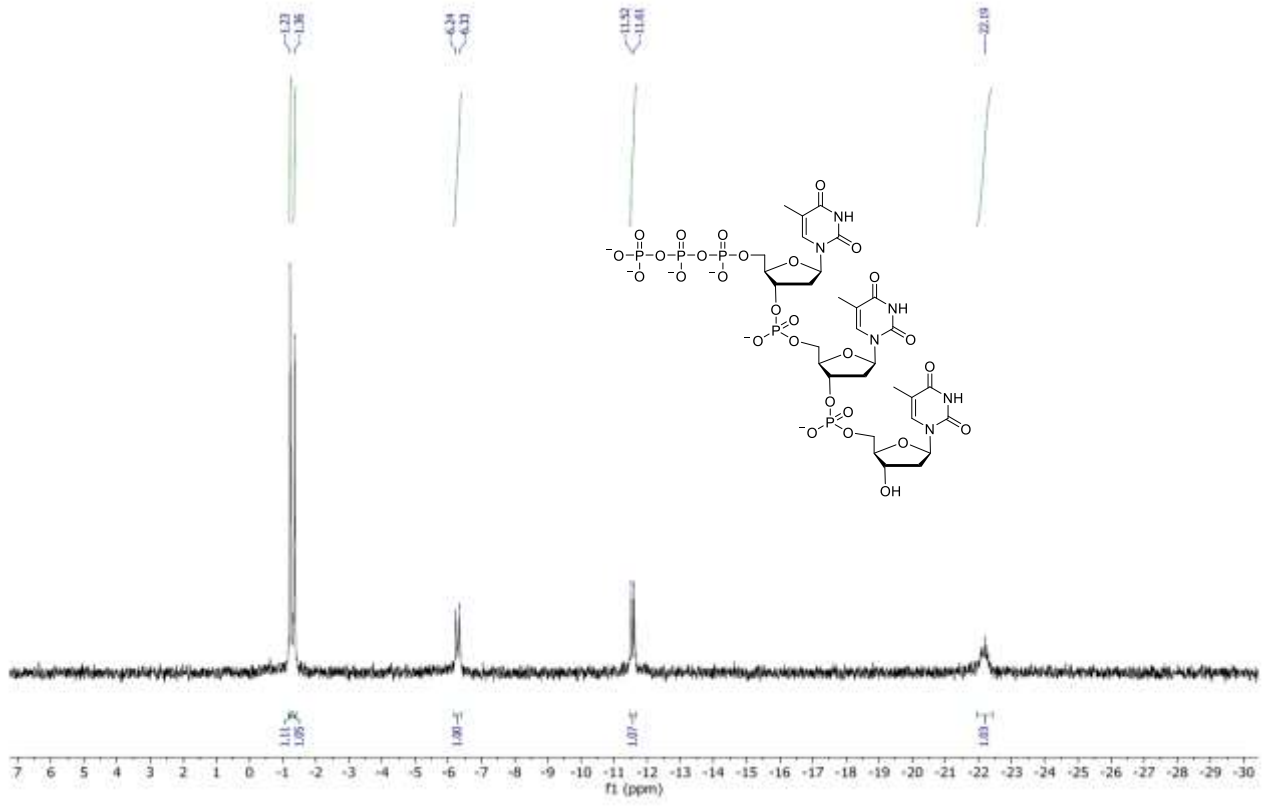
Trinucleotide triphosphate **4c** (8.4 mg, 8 µmol, 15%) was prepared according to the general procedure. ^{1}H NMR (500 MHz, D_2O) 7.66 – 7.60 (m, 3H), 6.19 – 6.14 (m, 3H), 4.98 – 4.92 (m, 2H), 4.46 – 4.45 (m, 1H), 4.29 – 4.26 (m, 2H), 4.08 – 4.02 (m, 7H), 2.43 – 2.38 (m, 3H), 2.25 – 2.18 (m, 3H), 1.81 – 1.78 (m, 9H), ^{31}P NMR (203 MHz, D_2O) δ 55.21 (m,1P); 54.96 (m, 1P), 0.45 (d, *J* = 20.4 Hz, 1P), -11.87 (d, *J* = 18.9 Hz, 1P), -22.84 (t, *J* = 20.1 Hz, 1P). HRMS calcd. for $C_{30}H_{41}N_6O_{26}P_5S_2$ $[M\text{-}2H]^{2-}$: 1120.0211, found: 1120.0208.

2.2 Copies of NMR Spectra

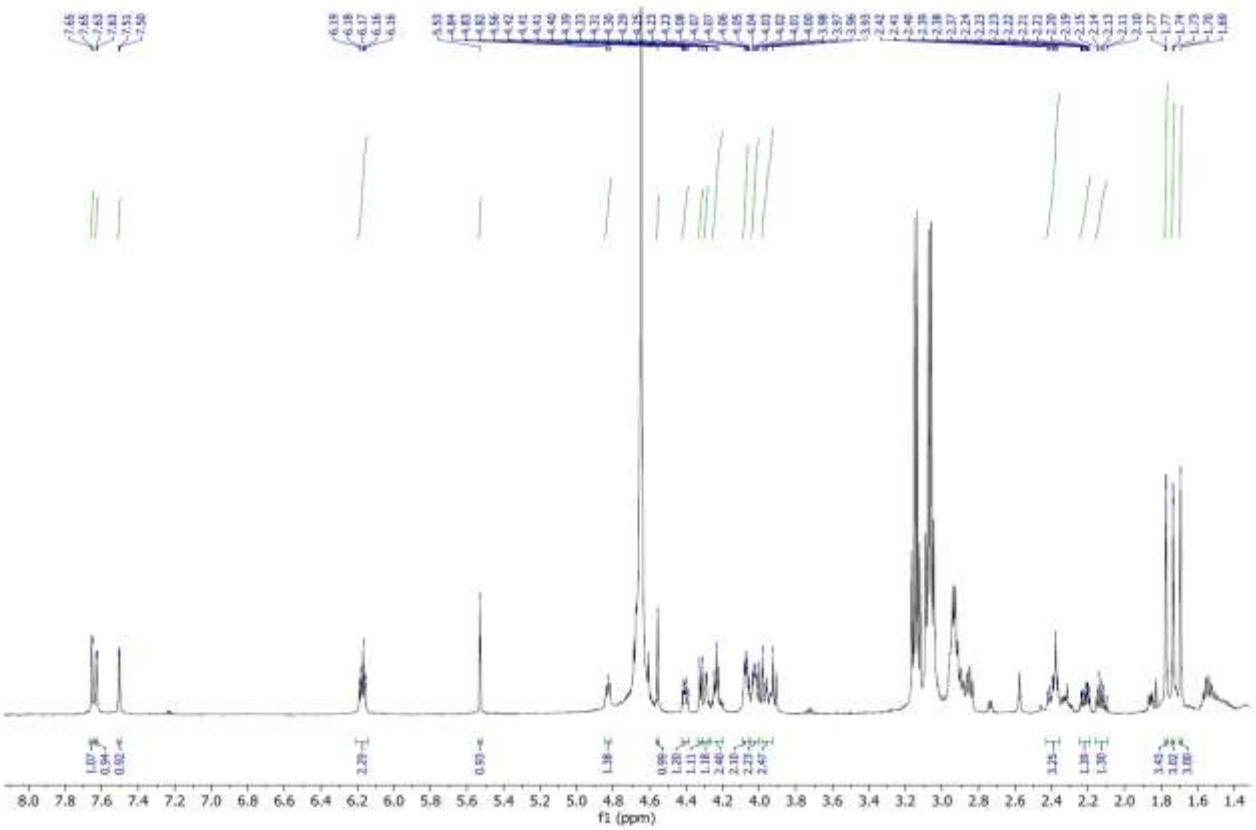
^{1}H NMR (D_2O, 500 MHz) spectrum of **Trinucleotide-5′-*O*-triphosphate 4a**



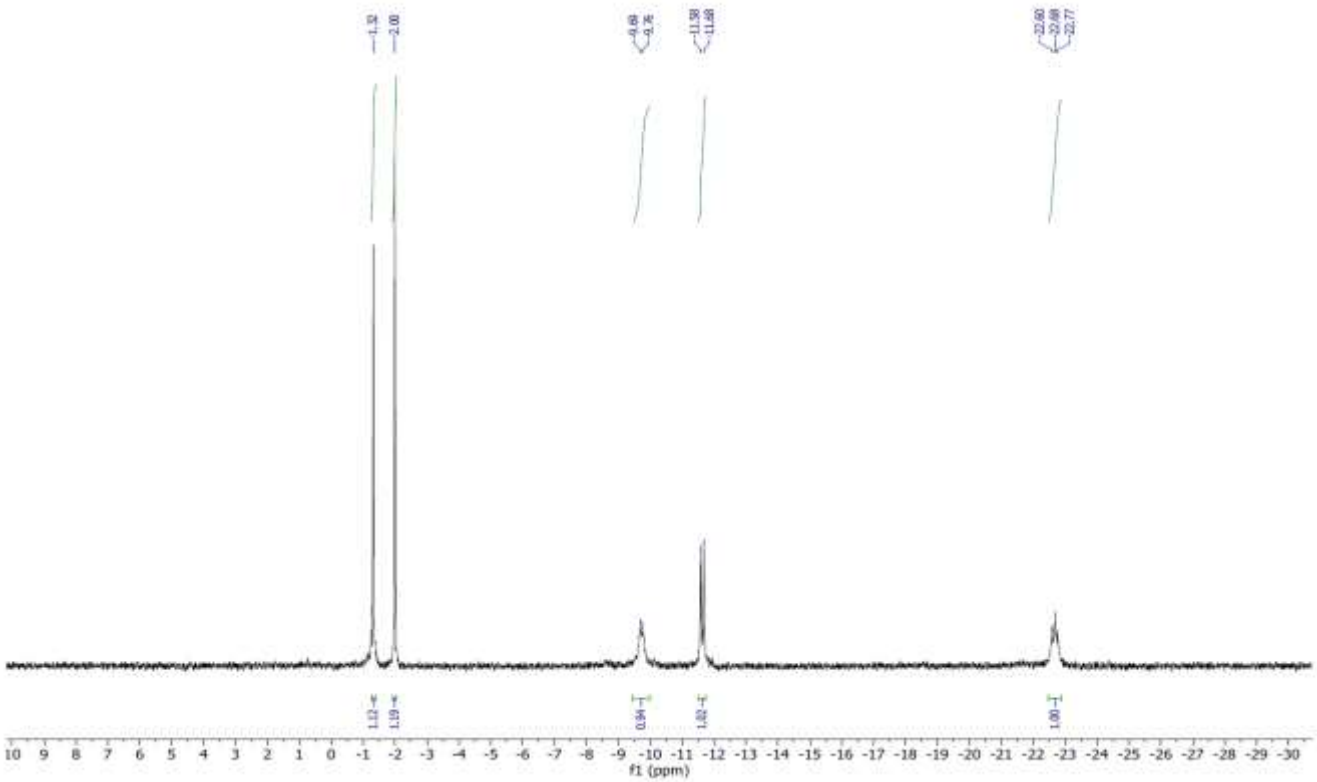
^{31}P NMR (D_2O, 203 MHz) spectrum of **Trinucleotide-5′-*O*-triphosphate 4a**



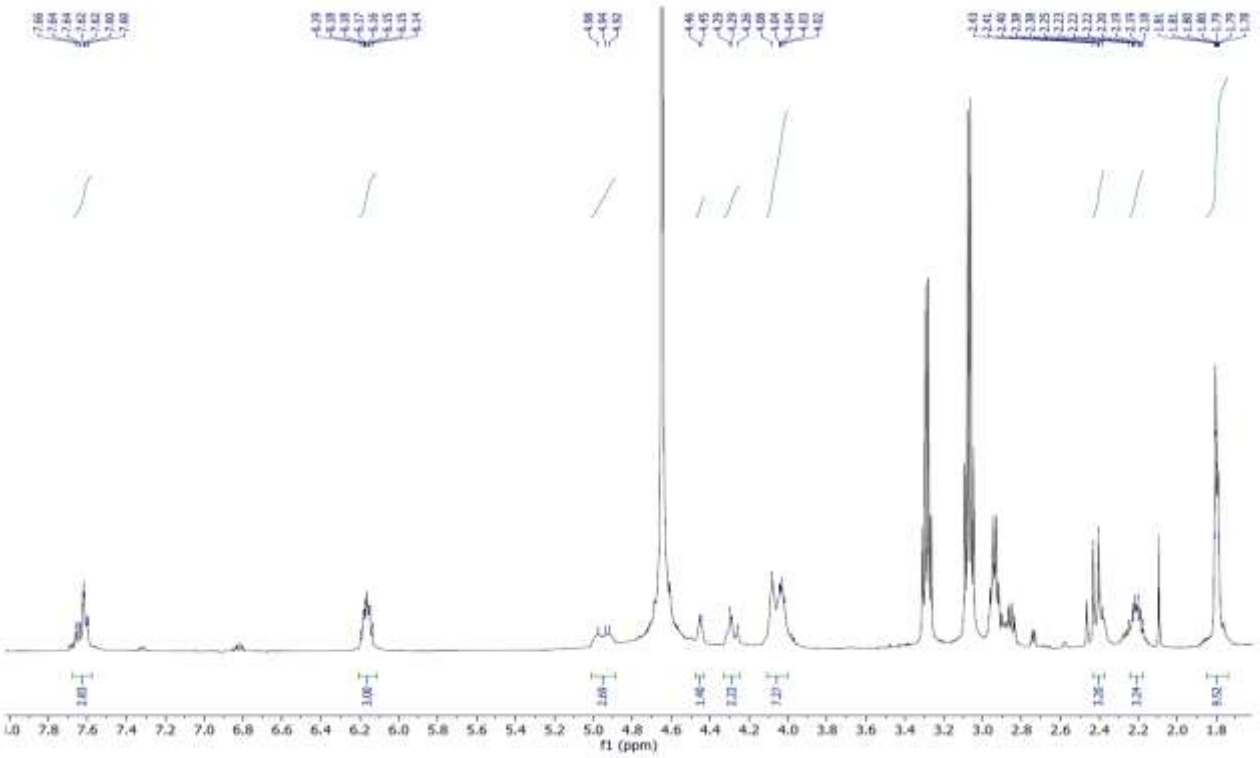
^{1}H NMR (D_2O, 500 MHz) spectrum of **LNA-trinucleotide-5′-*O*-triphosphate 4b**



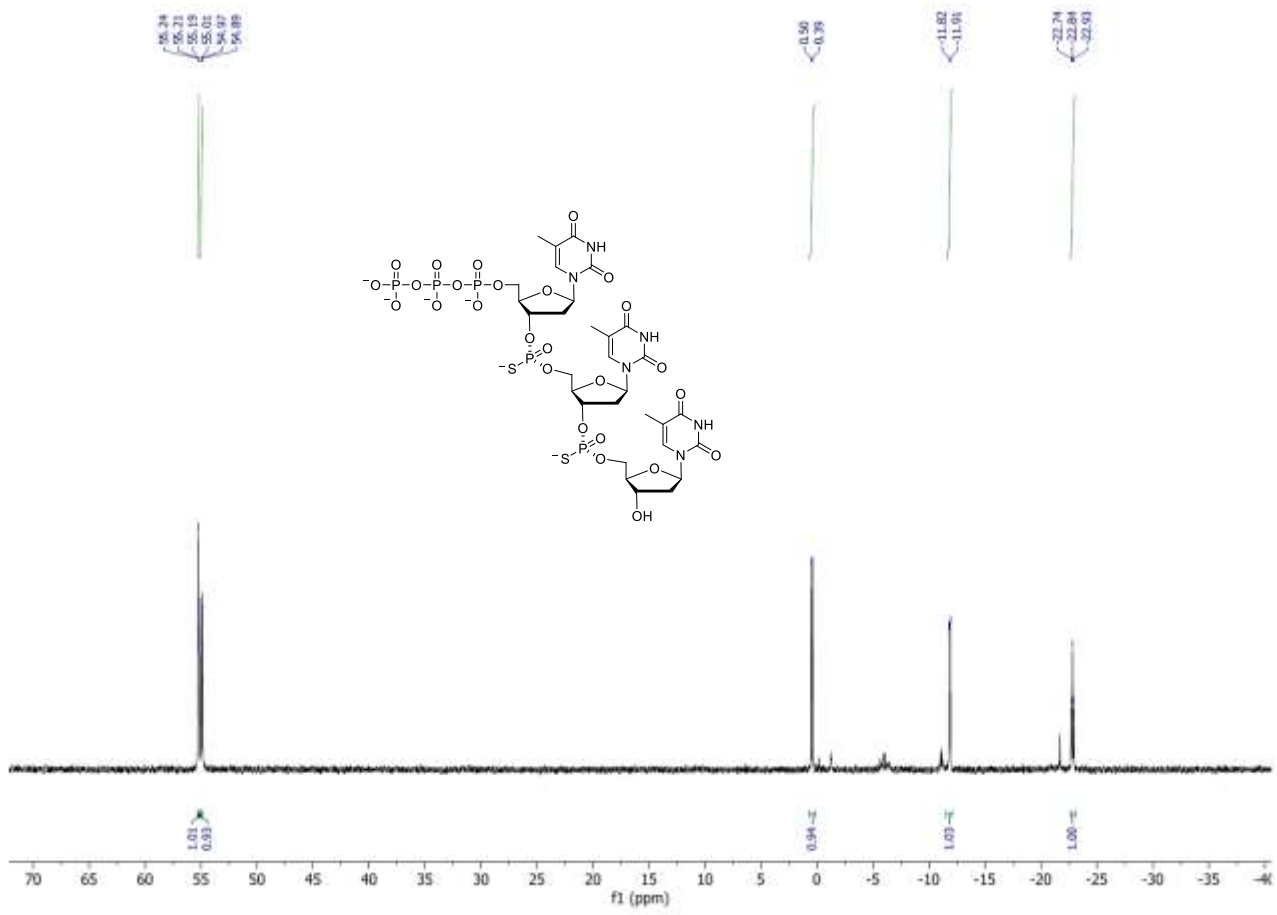
^{31}P NMR (D_2O, 203 MHz) spectrum of **LNA-trinucleotide-5′-*O*-triphosphate 4b**



^{1}H NMR (D_2O, 500 MHz) spectrum of **Thioate-trinucleotide-5′-*O*-triphosphate 4c**



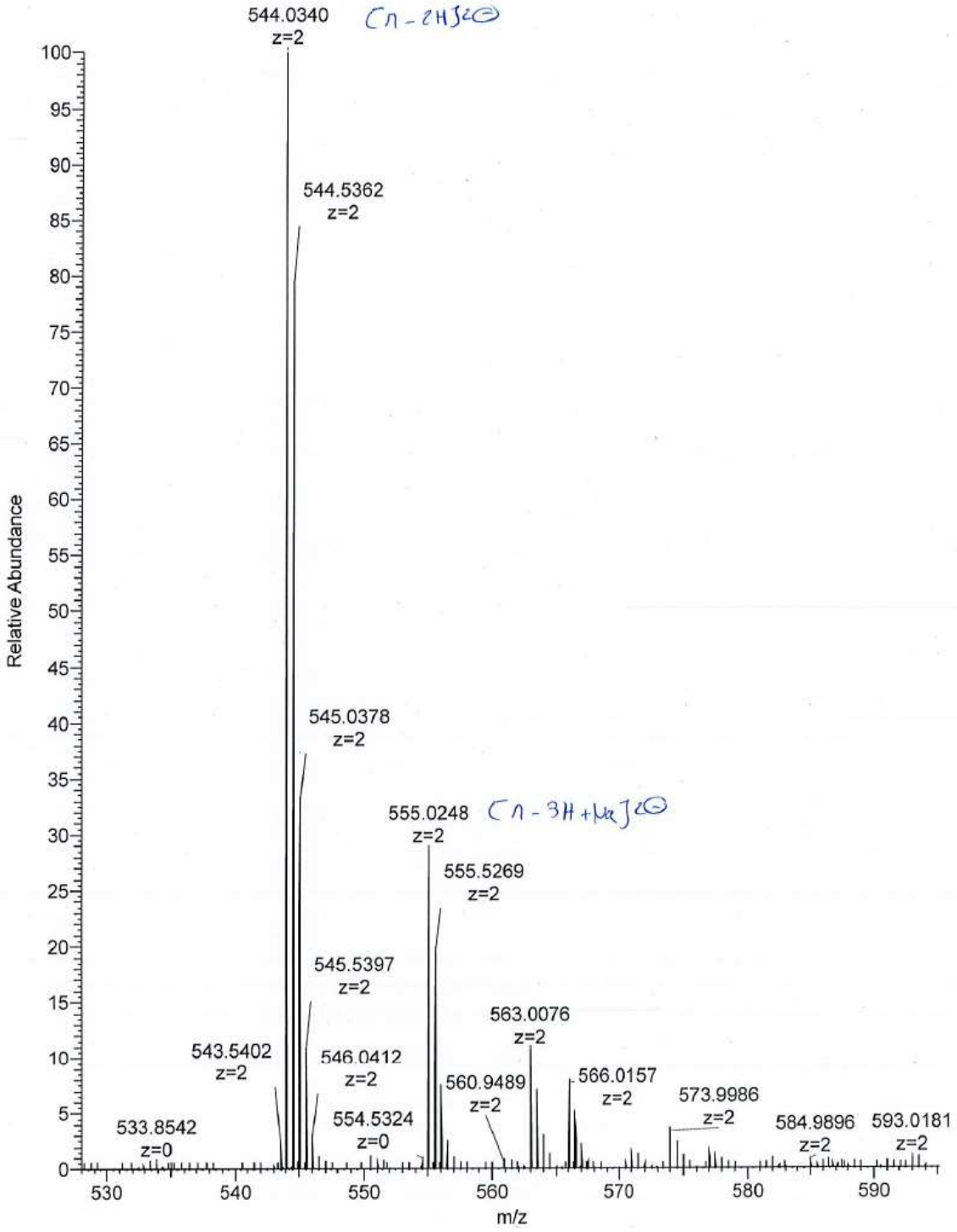
^{31}P NMR (D_2O, 203 MHz) spectrum of **Thioate-trinucleotide-5′-*O*-triphosphate 4c**



2.3 Copies of HRMS Spectra

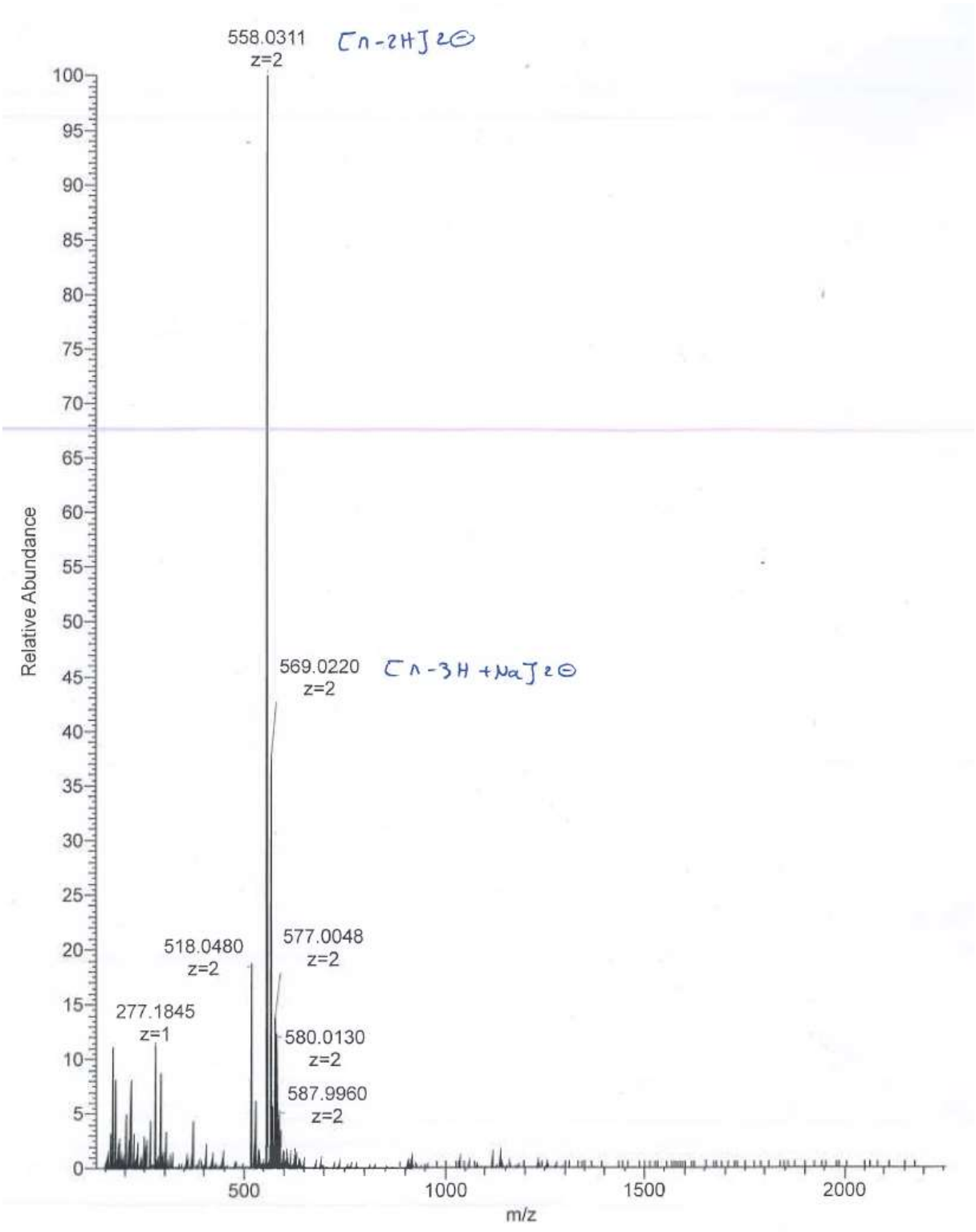
Trinucleotide-5′-*O*-triphosphate 4a

HRMS calcd. for $C_{30}H_{41}N_6O_{28}P_5$ $[M-2H]^{2-}$: 1088.0668, found: 1088.0680.



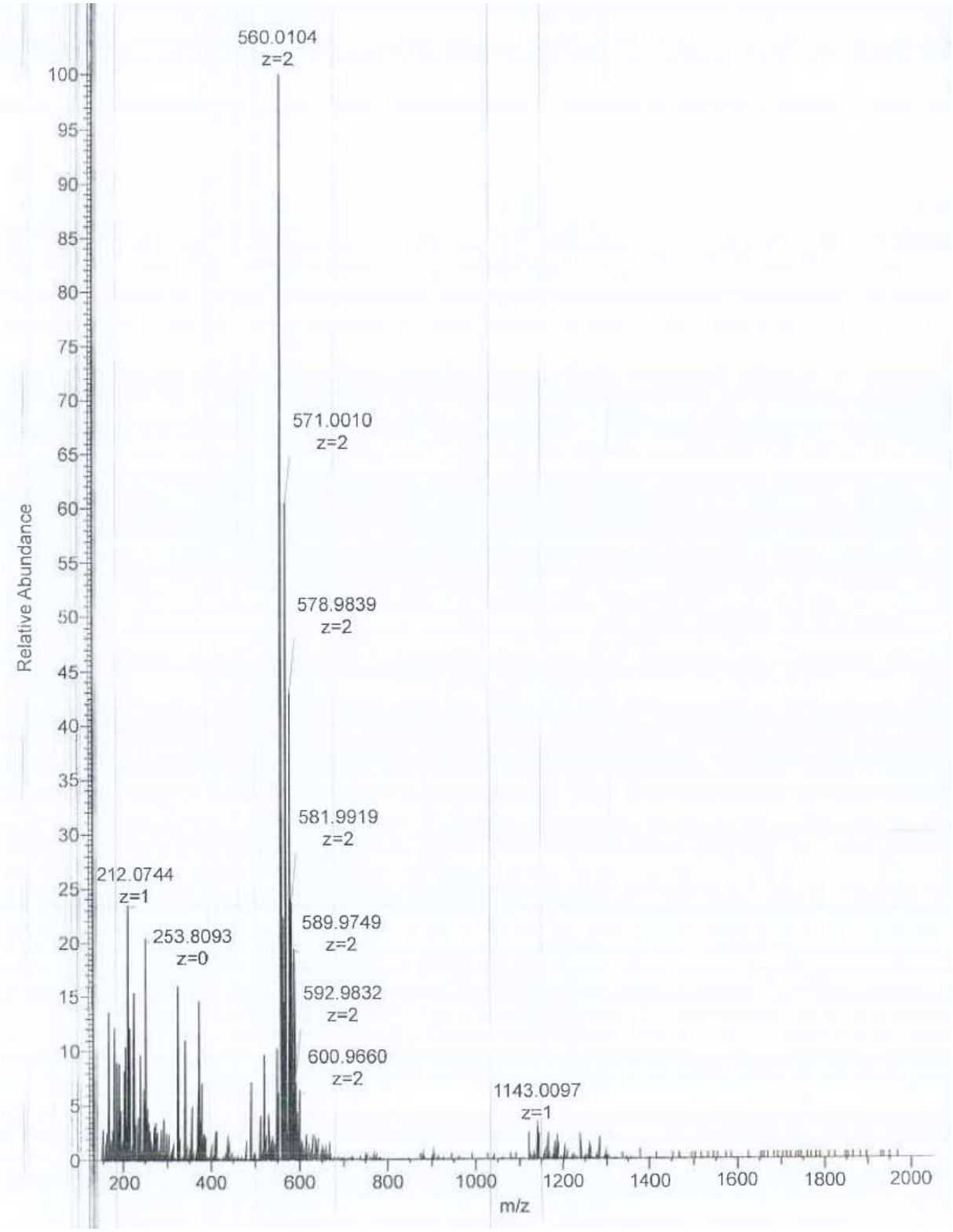
LNA-trinucleotide-5′-*O*-triphosphate 4b

HRMS calcd. for $C_{31}H_{41}N_6O_{29}P_5$ [M-2H]$^{2-}$: 1116.0617, found: 1116.0622.



Phosphorothioate trinucleotide-5′-*O*-triphosphate 4c

HRMS calcd. for $C_{30}H_{41}N_6O_{26}P_5S_2$ $[M-2H]^{2-}$: 1120.0211, found: 1120.0208.



3. Enzymatic reactions

3.1 Template-independent DNA extension reactions (TdT, PUP)

TdT reactions of DNA primer P1 with dT_3TPs 4a-c

DNA primer **P1**: 5'- FAM-TAC GAC TCA CTA TAG CCT C -3' (19 nt); MW: 6244.

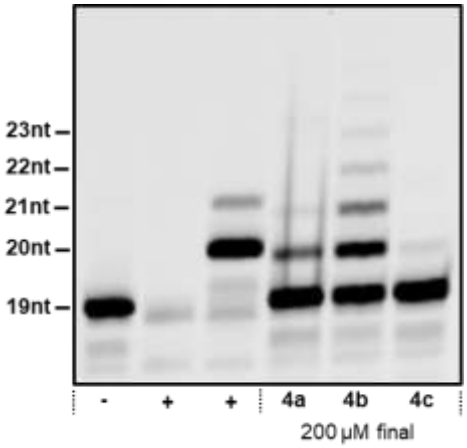


Figure SI2. Gel image (PAGE 20%) for analysis of TdT reactions. (**-**) – negative control in the absence of TdT enzyme; first from the left (**+**) – positive control using dTTP; second from the left (**+**) – positive control using (LNA-T)TP; (**4a-c)** – dT_3TPs (200 µM final).

Analytical scale reaction conditions (10 µL): DNA Primer **P1** 10µM (1 µL), dT_3TPs **4a-c** 1mM (2 µL), TdT buffer 10X (1 µL), Water (4 µL), TdT enzyme (0.5 µL), TIPP (0.5 µL), $MnCl_2$ 10 mM (1 µL). Reaction mixtures were incubated at 37 °C for 3 hours.

PUP reactions of RNA primer P2 with dT_3TPs 4a-c

RNA primer **P2**: 5'- FAM-rCrArG rUrCrG rGrArU rCrGrC rArGrU rCrArG (18 nt), MW: 6308.

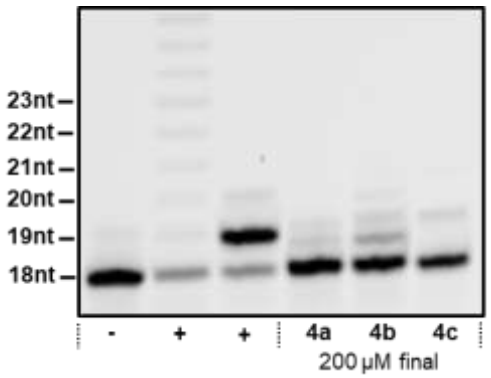


Figure SI3. Gel image (PAGE 20%) for analysis of PUP reactions. (**-**) – negative control in the absence of PUP enzyme; first from the left (**+**) – positive control using UTP; second from the left (**+**) – positive control using (LNA-T)TP; (**4a-c)** – dT_3TPs (200 µM final).

Analytical scale reaction conditions (10 µL): RNA Primer **P2** 10µM (2 µL), dT_3TPs **4a-c** 1mM (2 µL), NEBuffer 2 10X (1 µL), Water (3 µL), PolyU Polymerase (0.5 µL), RNAse Inhibitor Murine (0.5 µL), $MnCl_2$ 10 mM (1 µL). Reaction mixtures were incubated at 37 °C for 3 hours.

3.2 Template-dependent PEX reactions

PEX reactions of DNA primer P3 on DNA templates T1-T2 with dT_3TPs 4a-c

DNA primer **P3**: 5'- FAM-CAT GGG CGG CAT GGG -3' (15 nt); MW: 5211.

DNA Template **T1**: 5'- $(\underline{AAA})_4$ CCC ATG CCG CCC ATG -3' (27 nt); MW: 8228.

DNA Template **T2**: 5'- GTC AAA CCC TGG AAA CGT AAA CGC AAA GG AAA CCC ATG CCG CCC ATG -3' (47 nt); MW: 14405.

Analytical scale general reaction conditions (10 µL): DNA Primer **P3** 10µM (1 µL), DNA Template **T1-T2** 10µM (1.5 µL), Water (0.5 µL) were mixed and hybridized. Next, other components of the reaction were added depending on the DNA polymerase used:

- **a)** **HemoKlen Taq** polymerase (1 µL), Hemo KlenTaq buffer 5X (2 µL), dT_3TPs **4a-c** 1mM (2 µL), Water (2 µL). Reaction mixtures were incubated at 60 °C for 30 min;
- **b)** **Bst 2.0** polymerase (1 µL), Isothermal buffer 10X (1 µL), dT_3TPs **4a-c** 1mM (2 µL), Water (2 µL). Reaction mixtures were incubated at 60 °C for 30 min;
- **c)** **Vent (*exo*$^-$)** polymerase (1 µL), Thermopol buffer 10X (1 µL), dT_3TPs **4a-c** 1mM (2 µL), Water (2 µL). Reaction mixtures were incubated at 60 °C for 30 min;
- **d)** ***Taq*** polymerase (1 µL), Standard *Taq* buffer 10X (1 µL), dT_3TPs **4a-c** 1mM (2 µL), Water (2 µL). Reaction mixtures were incubated at 60 °C for 30 min;
- **e)** **Sulfolobus** polymerase (1 µL), Thermopol buffer 10X (1 µL), dT_3TPs **4a-c** 1mM (2 µL), Water (2 µL). Reaction mixtures were incubated at 55 °C for 30 min;
- **f)** **Deep Vent (*exo*$^-$)** polymerase (1 µL), Thermopol buffer 10X (1 µL), dT_3TPs **4a-c** 1mM (2 µL), Water (2 µL). Reaction mixtures were incubated at 55 °C for 30 min;
- **g)** **Klenow (*exo*$^-$)** polymerase (1 µL), NEBuffer 2 10X (1 µL), dT_3TPs **4a-c** or **dTTP** or **α-thio-TTP** 1mM (2 µL), Water (2 µL). Reaction mixtures were incubated at 37 °C for 30 min.

Preparative scale general reaction conditions (100 µL): DNA Primer **P3** 100µM (5 µL), DNA Template **T1-T2** 100µM (7.5 µL), Water (5.0 µL) were mixed and hybridized. Next, other components of the reaction were added depending on the DNA polymerase used:

- **a)** ***Taq*** polymerase (20 µL), Standard *Taq* buffer 10X (10 µL), dT_3TPs **4a-c** 5mM (20 µL), Water (32.5 µL). Reaction mixtures were incubated at 60 °C for 30 min;
- **b)** **Klenow (*exo*$^-$)** polymerase (20 µL), NEBuffer 2 10X (10 µL), dT_3TPs **4a-c** or **dTTP** or **α-thio-TTP** 5mM (20 µL), Water (32.5 µL). Reaction mixtures were incubated at 37 °C for 30 min.

After incubation, the reaction mixture was diluted with water to 500 µL final volume. Then washed and concentrated using an Amicon Ultra centrifugal filter (10kDa). The resulting DNA concentrate was used for further LCMS analysis.

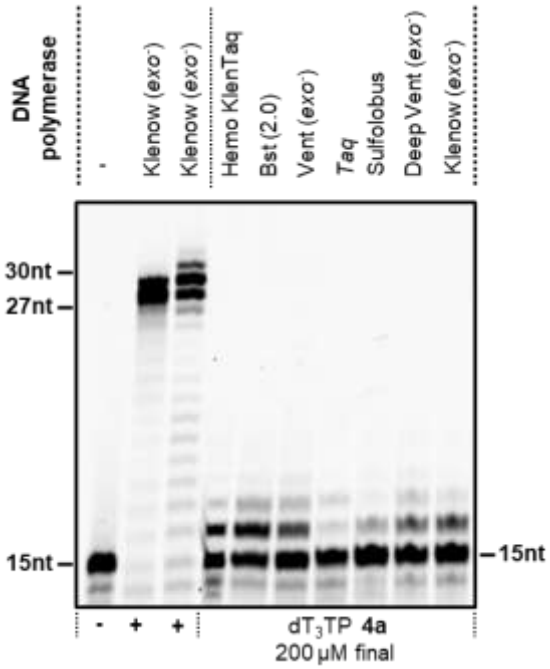


Figure SI4A. Gel image (PAGE 20%) for analysis of PEX reactions of DNA primer **P3** on DNA template **T1** with dT_3TP **4a** (200 µM final) over 30 minutes. (**-**) – negative control in the absence of DNA polymerase; first from the left (**+**) – positive control using dTTP (Klenow *exo*$^-$); second from the left (**+**) – positive control using α-thio-TTP (Klenow *exo*$^-$).

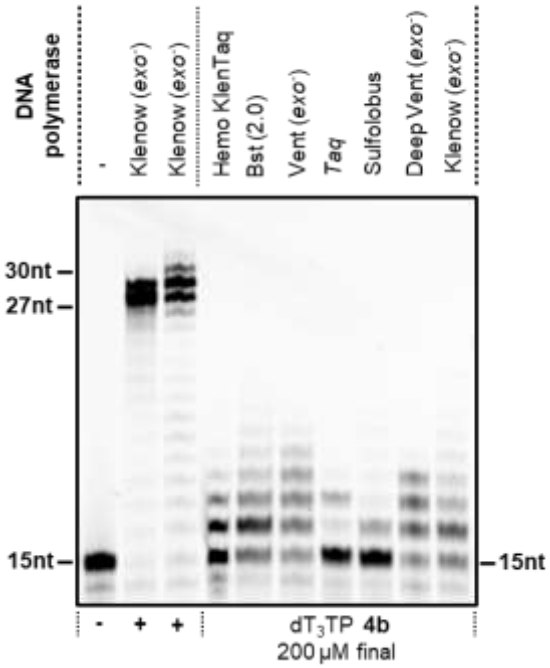


Figure SI4B. Gel image (PAGE 20%) for analysis of PEX reactions of DNA primer **P3** on DNA template **T1** with dT_3TP **4b** (200 µM final) over 30 minutes. (**-**) – negative control in the absence of DNA polymerase; first from the left (**+**) – positive control using dTTP (Klenow *exo*$^-$); second from the left (**+**) – positive control using α-thio-TTP (Klenow *exo*$^-$).

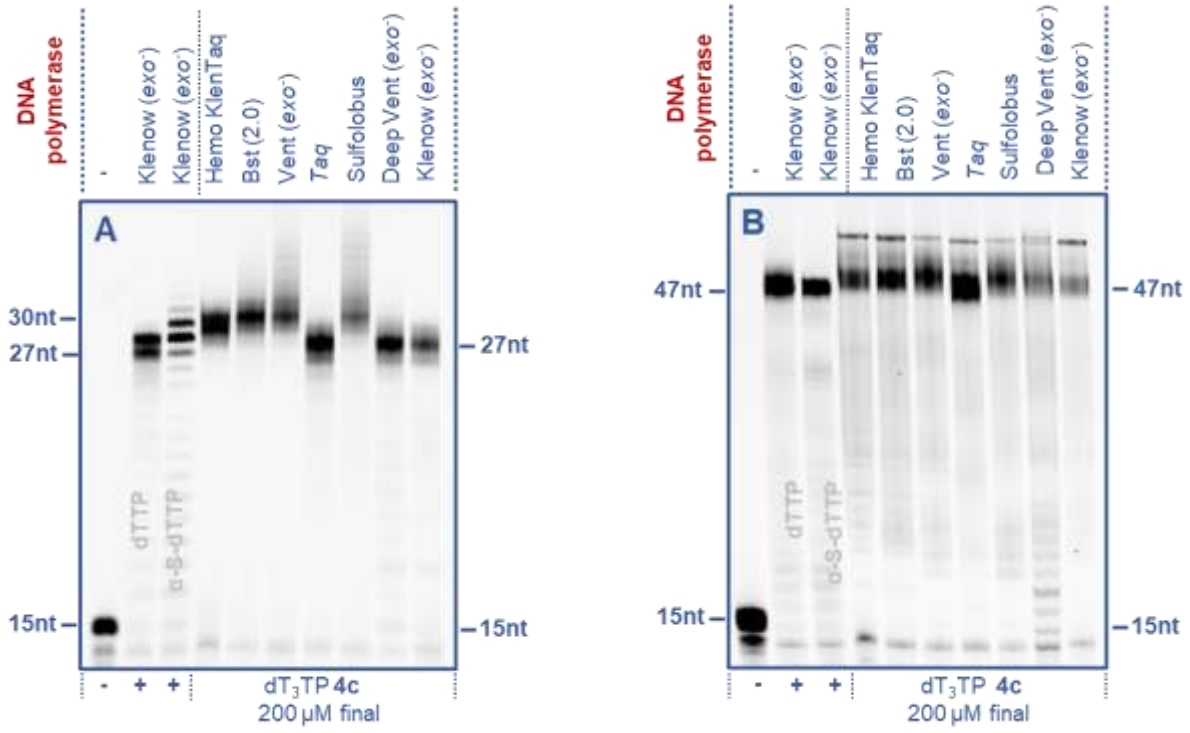


Figure 3A-B. Gel image (PAGE 20%) for analysis of PEX reactions of DNA primer **P3** on DNA Templates **T1-T2** with dT_3TP **4c** (200 µM final) over 30 minutes. (**-**) – negative control in the absence of DNA polymerase; first from the left (**+**) – positive control using dTTP (Klenow *exo*$^-$); second from the left (**+**) – positive control using α-thio-TTP (Klenow *exo*$^-$).

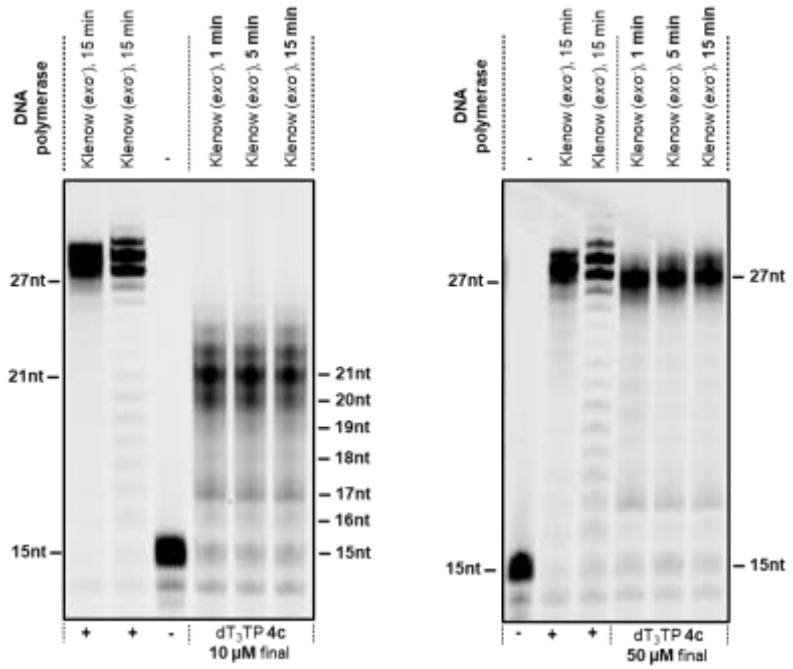


Figure SI5. Gel image (PAGE 20%) for analysis of PEX reactions of DNA primer **P3** on DNA Template **T1** with dT_3TP **4c** (10 and 50 µM final) over 1-15 min timeframe. (**-**) – negative control in the absence of DNA polymerase; first from the left (**+**) – positive control using dTTP (Klenow *exo*$^-$); second from the left (**+**) – positive control using α-thio-TTP (Klenow *exo*$^-$).

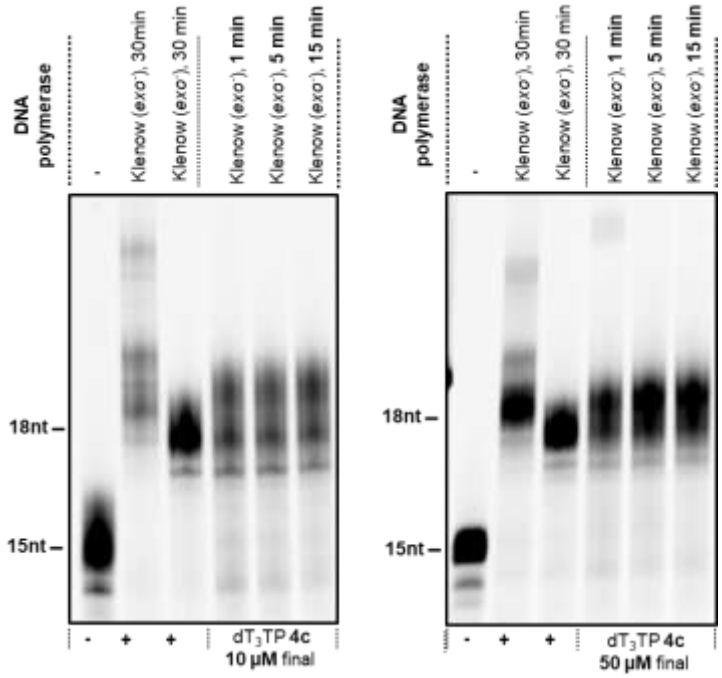


Figure SI6. Gel image (PAGE 20%) for analysis of PEX reactions of DNA primer **P3** on DNA Template **T2** with dT_3TP **4c** (10 and 50 µM final) over 1-15 min timeframe. (**-**) – negative control in the absence of DNA polymerase; first from the left (**+**) – positive control using dTTP (Klenow *exo*$^-$); second from the left (**+**) – positive control using α-thio-TTP (Klenow *exo*$^-$).

3.3 Evaluation of stability of trinucleotide triphosphate dT$_3$TP 4c

We first analyzed pure dTTP and dT$_3$TP **4c** by running analytical ion exchange HPLC (using an Äkta™ pure system GE Healthcare equipped with Thermo Scientific™ DNAPac™ PA100 column, Buffer A: 10 mM TEAB to 100% Buffer B: 1 M TEAB) (Figures SI7-8). We next incubated a 10 µL of a 1 mM solution of dT$_3$TP **4c** in the presence of 5 µL of the (Klenow *exo*$^-$) DNA polymerase and 5 µL of NEB2 buffer at 37°C for either 5 min or 1 hour. Then we analyzed directly the crude rection mixtures by analytical ion exchange HPLC (Figures SI9-10). To conclude, under these reaction conditions, we did not observe the formation of dTTP or other degradation products suggesting that dT$_3$TP **4c** is not readily degraded.

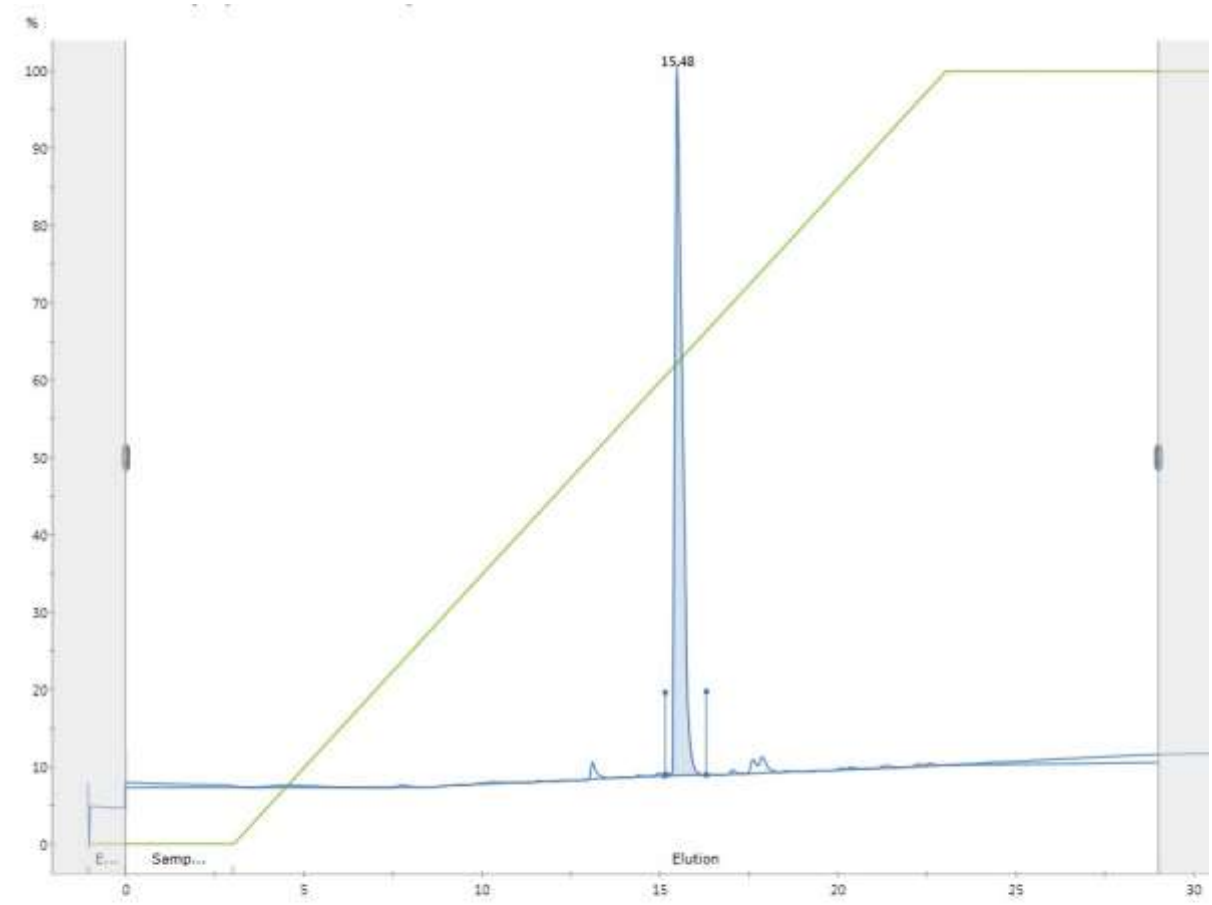


Figure SI7. HPLC trace of dTTP. Peak A (100%), RT (15.5 min). (Y-Axis): Buffer A (10 mM TEAB) to 100% Buffer B (1 M TEAB).

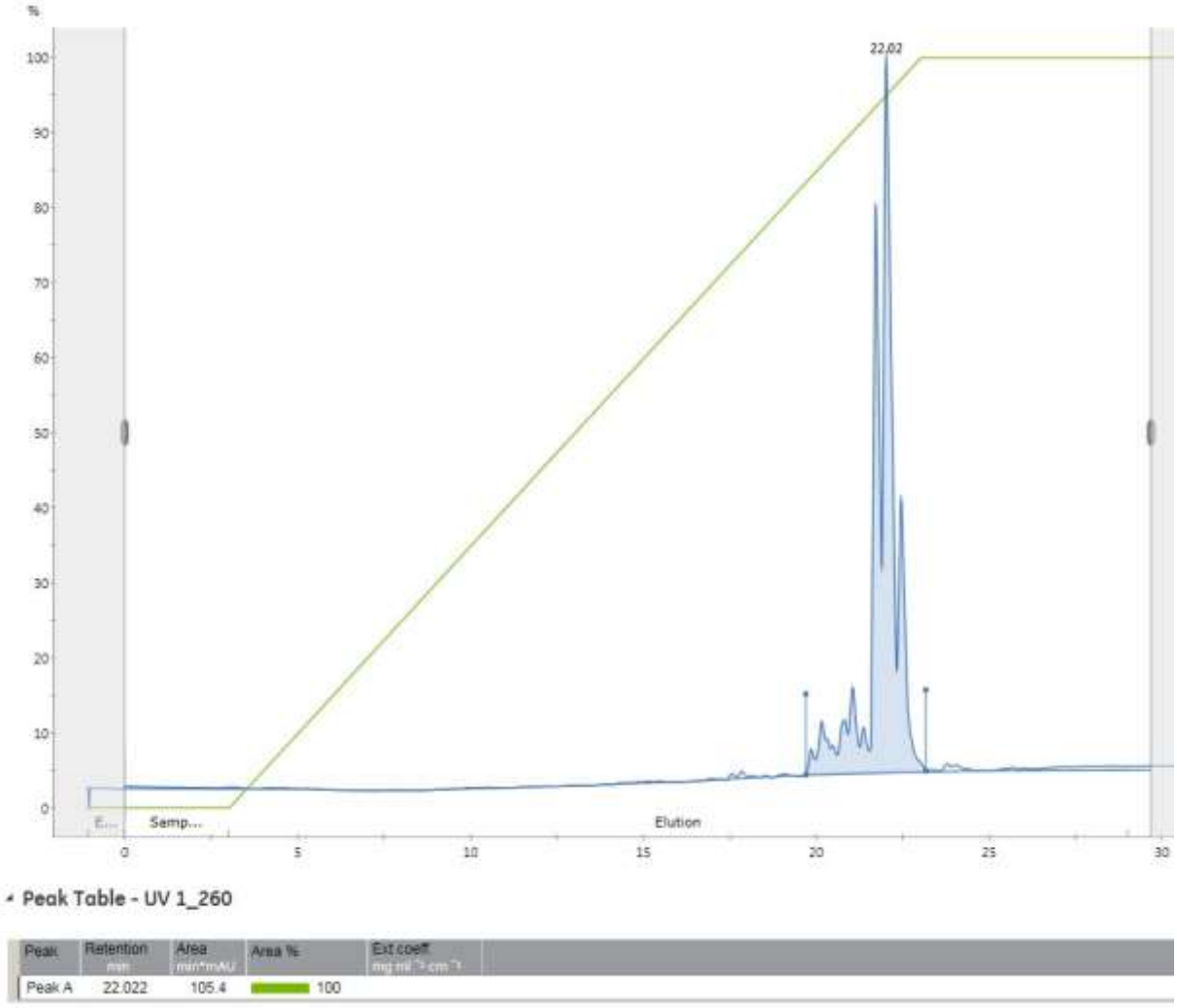


Figure SI8. HPLC trace of dT_3TP **4c** as a mixture of Rp and Sp stereoisomers. Peak A (100%), RT (22.0 min). (Y-Axis): Buffer A (10 mM TEAB) to 100% Buffer B (1 M TEAB).

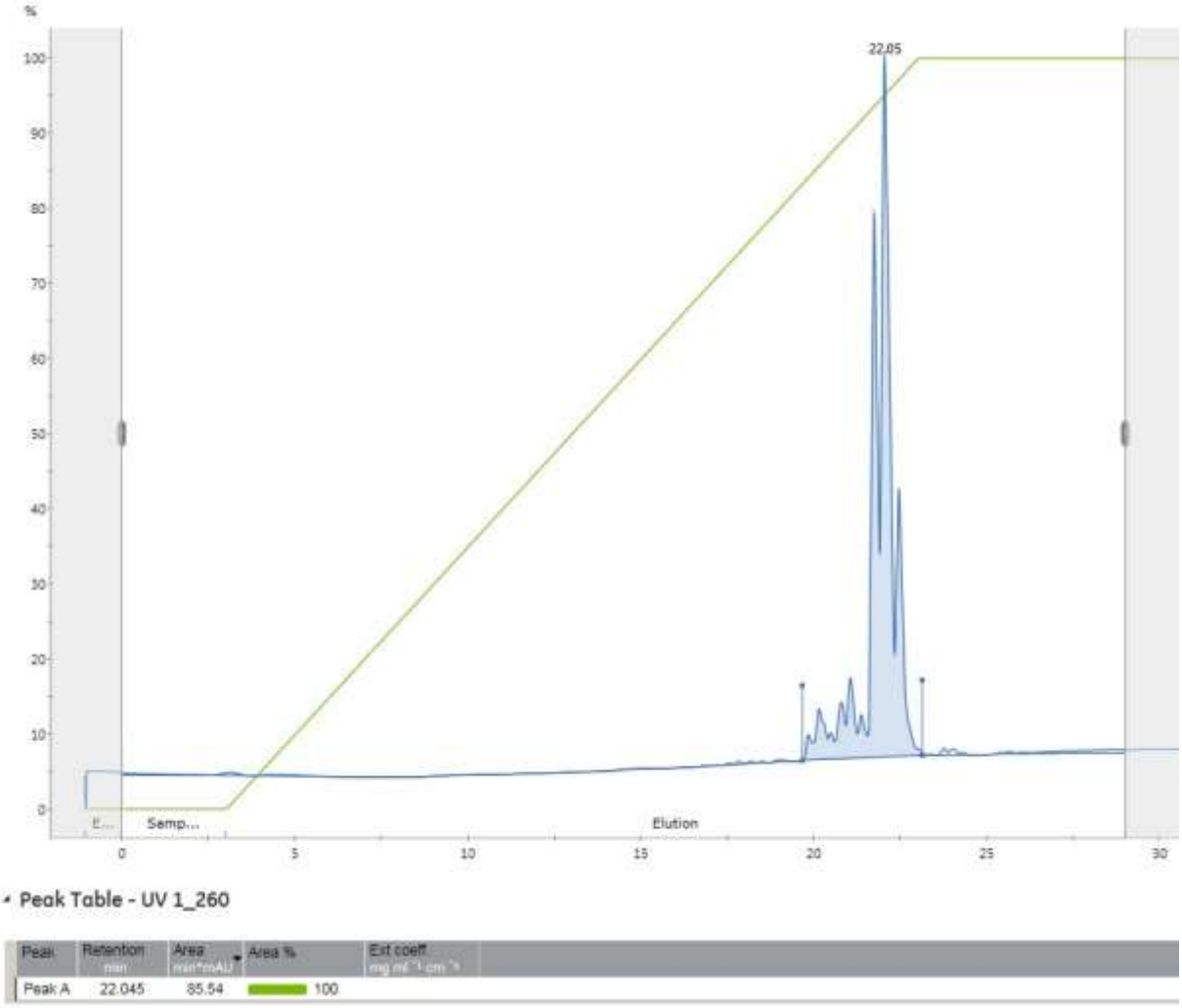
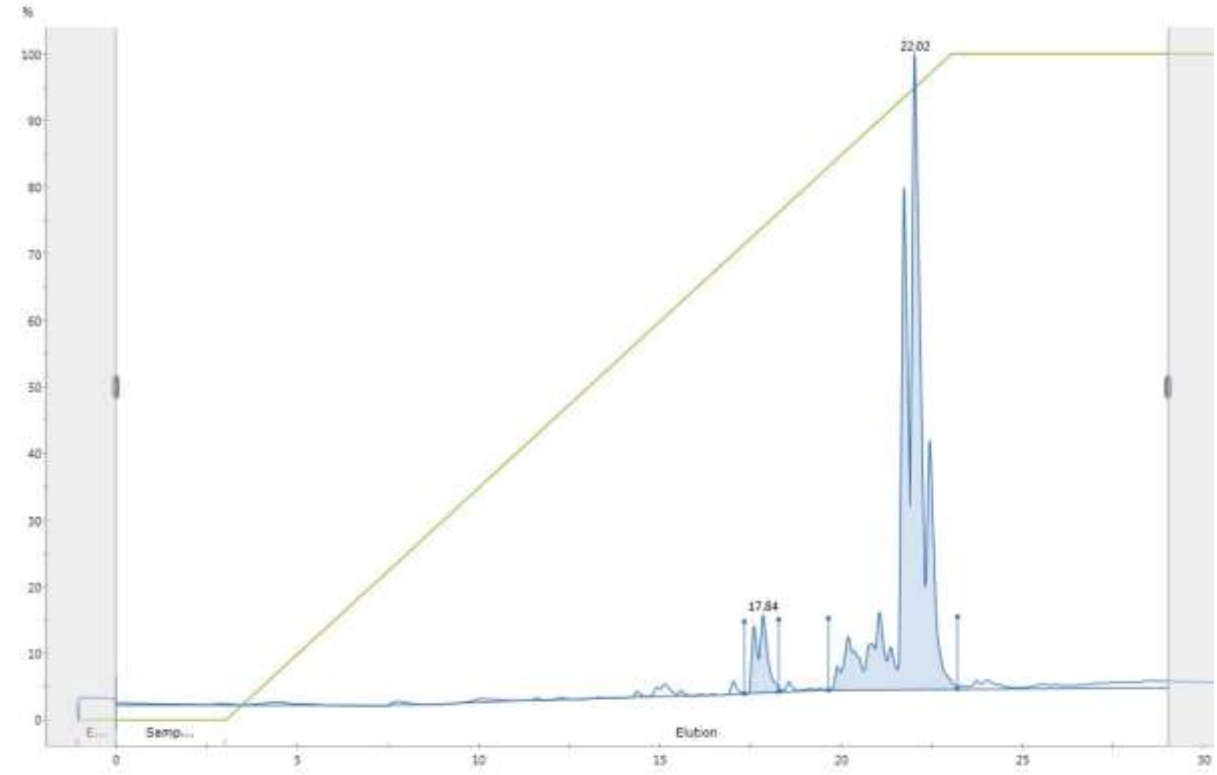


Figure SI9. HPLC trace of crude reaction mixture (*5 minutes*) of dT_3TP **4c**, (Klenow *exo*$^-$) polymerase and NEB2 buffer. Peak A (100%), RT (22.0 min). (Y-Axis): Buffer A (10 mM TEAB) to 100% Buffer B (1 M TEAB).



Peak Table - UV 1_260

Peak	Retention	Area	Area %	Ext coeff
Peak A	17.835	8.334	7.27	
Peak B	22.023	106.3	92.73	

Figure SI10. HPLC trace of crude reaction mixture (*1 hour*) of dT_3TP **4c**, (Klenow *exo*$^-$) polymerase and NEB2 buffer. Peak A (93%), RT (22.0 min). (Y-Axis): Buffer A (10 mM TEAB) to 100% Buffer B (1 M TEAB).

4. LCMS analyses of enzymatic reactions

LCMS settings and method

Chromatographic separations for LC–MS experiments were performed on a Thermo Scientific™ Vanquish™ Flex Binary UHPLC system (Thermo Fisher Scientific, Reinach, Switzerland).

The column used for all separations was a Waters Aquity Premier BEH C18 Peptide 2.1*50 mm 1.7 µm 300A (Waters (CH) AG, Baden-Dättwil, Switzerland). The solvents were A: 15 mM Amylamine (ALDRICH, W424201 (in-house re-distilled), CAS 110-58-7, SIGMA-ALDRICH CHEMIE GMBH (CH), Buchs Switzerland) and 50 mM 1,1,1,3,3,3-HEXAFLUORO-2-PROPANOL (HFIP, ACROS ORGANICS, ACR14754 (99.5+%, PURE), CAS 920-66-1, ThermoFisher Scientific, Reinach Switzerland) in water (Milli-Q® IQ 7000, Millipore, Merck & Cie, Schafhausen, Switzerland) and B: methanol/acetonitrile (9/1;v/v) both gradient grade. Both solvents contained 1 µM Ethylenediaminetetraacetic acid disodium salt dihydrate (EDTA), SIGMA, E5134-100G (99.0-101.0%, titration)), CAS 6381-92-6, SIGMA-ALDRICH CHEMIE GMBH (CH), Buchs Switzerland helping to suppress metal adducts forming in the mass spectrometer. A flow rate of 400 µL min^{-1} was applied, and the column compartment was held at 80°C. The gradient system was A vs B starting at 5%B with a hold for 0.2 min before a linear raise to 45%B at.5 min followed by a 0.5 min step to 100% B for flushing the column for 1 min before coming back to the initial 5% B for reinjection.

For mass spectrometric data acquisition, a Thermo Scientific™ Fusion Lumos™ Hybrid Quadrupole-Orbitrap mass spectrometer equipped with a heated electrospray ionization-II (HESI-II) probe in a standard Thermo Scientific™ Ion Max™ ion source (Thermo Fisher Scientific, San José, CA, USA) was used. Data acquisition was performed with Thermo Scientific™ Xcalibur 4.5. HR LC–MS measurements were performed under Thermo Scientific™ Xcalibur™ Orbitrap Fusion Lumos Tune Application 3.5. The mass range was 600-2000 Da at a resolution of 120K. In addition, the full DAD but also a 260 nm UV trace were acquired.

MS raw files were exported to ThermoFisher Scientific BioPharma Finder (BPF) 5.1 software and analyzed using the Intact Mass Analysis deconvolution feature either against the full structure of the compounds or just against the molecular mass. All data evaluation was done based on the Thermo Scientific BioPharma Finder User Guide Software Version 5.1, XCALI-98492 Revision A, July 2022.

Table SI1. Summary of PEX reactions of DNA primer P3 on DNA Template T1

Name	**m/z calc.**	**m/z found**	**Interpretation of the reaction products**
PEX of **dTTP** Klenow (exo-) (Positive control)	8857,4821 9161,5281 9465,5742 9769,6202	8857,5017 9161,5503 9465,5964 9769,6468	Mass of 12T additions (expected) Mass of 12T+1 additions Mass of 12T+2 additions Mass of 12T+3 additions
PEX of **4c** With *Taq*	8985,2994	8977,3241 8995,3220	Mass distribution resulted in two masses found (average 8986.3230).
PEX of **4c** With *Kf (exo-)*	8985,2994	8977,3019 8995,3158	Mass distribution resulted in two masses found (average 8986.3088).

Table SI2. Summary of PEX reactions of DNA primer P3 on DNA Template T2

Name	**m/z calc.**	**m/z found**	**Interpretation of the reaction products**
PEX of **dTTP** Klenow (exo-) (Positive control)	15035,4853 15916,6162 15940,6275 16229,6738	Not found 15916,5680 15940,5543 16229,6121	Mass of expected product Mass of expected product +CCT addition Mass of expected product +CTA addition Mass of expected product +CCTA addition
PEX of **4c** With *Taq*	15155,4853	15145,1574	Mass distribution resulted in a slightly lower mass
PEX of **4c** With *Kf (exo-)*	15155,4853 15458,2889	15145,1764 15458,2295	Mass distribution resulted in a slightly lower mass Mass of expected product +T addition

LCMS analysis chromatogram of PEX reaction of DNA primer **P3** on the DNA template **T1** with **dTTP** and **Klenow (*exo*$^-$)** polymerase (positive control reaction)

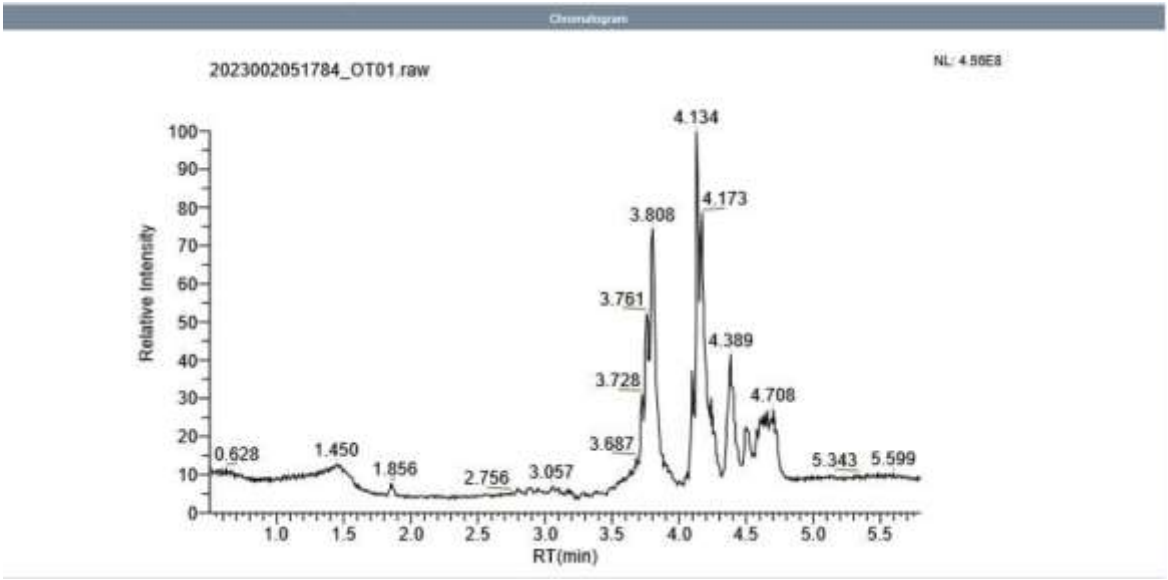


Figure SI11. LC chromatogram of PEX (**P3, T1, dTTP, Klenow**).

RT (3.68-3.81)

(*Row 6, 17*) - DNA template T1 (calc. 8228,4713; found 8203,4713; 8228,4637);

(*Row 4*) - DNA template T1 (+T addition) (calc. 8532,5173; found 8532,5261);

(*Row 2*) - DNA template T1 (+TT addition) (calc. 8836,5634; found 8836,5826).

RT (4.13-4.19)

(*Row 5*) - Expected DNA product (+one T addition) (calc. 9161,5281; 9161,5503);

(*Row 1*) - Expected DNA product (+two T additions) (calc. 9465,5742; found 9465,5964);

(*Row 3*) - Expected DNA product (+three T additions) (calc. 9769,6202; found 9769,6468).

Sliding Windows XTract Masses Table																	
Row Number	Sequence Name	Modification	Measured Spec Mass	Theoretic al Mass (Da)	Matched Mass Error (ppm)	Average Mass	Sum Intensity	Relative A bundance	Fractional Abundanc e	Number of Charge States	Charge State (M) Window	Number of Detected Isotopes	Delta Mass	Scan Range	Start Time (min)	Stop Time (min)	Apex RT
1			9465.596 40	0.00000	0.0	9470.21	14619482 8.04	100.00	14.98	9	5 - 13	21	0.00000	586 - 616	4.087	4.287	4.153
2			8836.582 56	0.00000	0.0	8840.89	10337460 6.02	70.71	10.57	9	5 - 13	16	-629.013 84	536 - 560	3.748	3.909	3.814
3			9769.646 75	0.00000	0.0	9774.41	85640317 .23	58.58	8.76	9	5 - 13	19	304.0503 6	592 - 620	4.127	4.314	4.187
4			8532.526 07	0.00000	0.0	8536.69	63766287 .71	43.62	6.52	8	5 - 12	17	-933.070 33	521 - 558	3.714	3.895	3.774
5			9161.550 31	0.00000	0.0	9166.02	48603982 .31	33.25	4.97	9	5 - 13	17	-304.046 68	580 - 617	4.046	4.227	4.114
6			8203.471 54	0.00000	0.0	8207.47	35432132 .69	24.24	3.62	8	5 - 12	14	-1262.12 486	527 - 548	3.687	3.828	3.741
7			1831.306 41	0.00000	0.0	1832.20	27655853 .70	18.91	2.77	3	1 - 3	11	-7634.28 998	263 - 278	1.813	1.933	1.862
8			10073.68 247	0.00000	0.0	10078.59	20434777 .61	13.98	2.09	8	6 - 13	12	608.0860 7	596 - 618	4.153	4.301	4.207
9			8856.582 09	0.00000	0.0	8864.90	16323043 .68	11.17	1.67	9	5 - 13	13	-615.014 91	542 - 562	3.788	3.922	3.841
10			10377.72 172	0.00000	0.0	10382.78	13318848 .26	9.11	1.36	9	6 - 14	16	912.1253 3	600 - 626	4.180	4.342	4.233
11			5975.892 33	0.00000	0.0	5978.36	13017688 .91	8.90	1.33	5	3 - 7	22	-4389.70 407	548 - 598	3.558	3.761	3.653
12			13114.14 943	0.00000	0.0	13120.54	12714201 .59	8.70	1.30	11	7 - 17	13	3648.552 64	623 - 644	4.335	4.478	4.382
13			3346.545 56	0.00000	0.0	3348.18	10959665 .26	7.50	1.12	4	2 - 5	14	-6119.05 083	431 - 451	3.015	3.155	3.064
14			10681.76 466	0.00000	0.0	10686.97	10891819 .06	7.45	1.11	9	6 - 14	15	1216.168 27	604 - 626	4.247	4.355	4.293
15			12810.09 179	0.00000	0.0	12816.34	10021923 .83	6.86	1.03	11	7 - 17	13	3344.499 91	620 - 640	4.314	4.449	4.375
16			3042.590 79	0.00000	0.0	3043.99	9266285 .70	6.29	0.94	4	2 - 5	14	-6423.09 560	405 - 426	2.833	2.986	2.896
17			8228.463 73	0.00000	0.0	8232.48	9162691 .31	6.27	0.94	7	5 - 11	11	-1237.13 266	525 - 543	3.674	3.794	3.728

LCMS analysis chromatogram of PEX reaction of DNA primer **P3** on the DNA template **T1** with dT_3TP **4c** and ***Taq*** polymerase

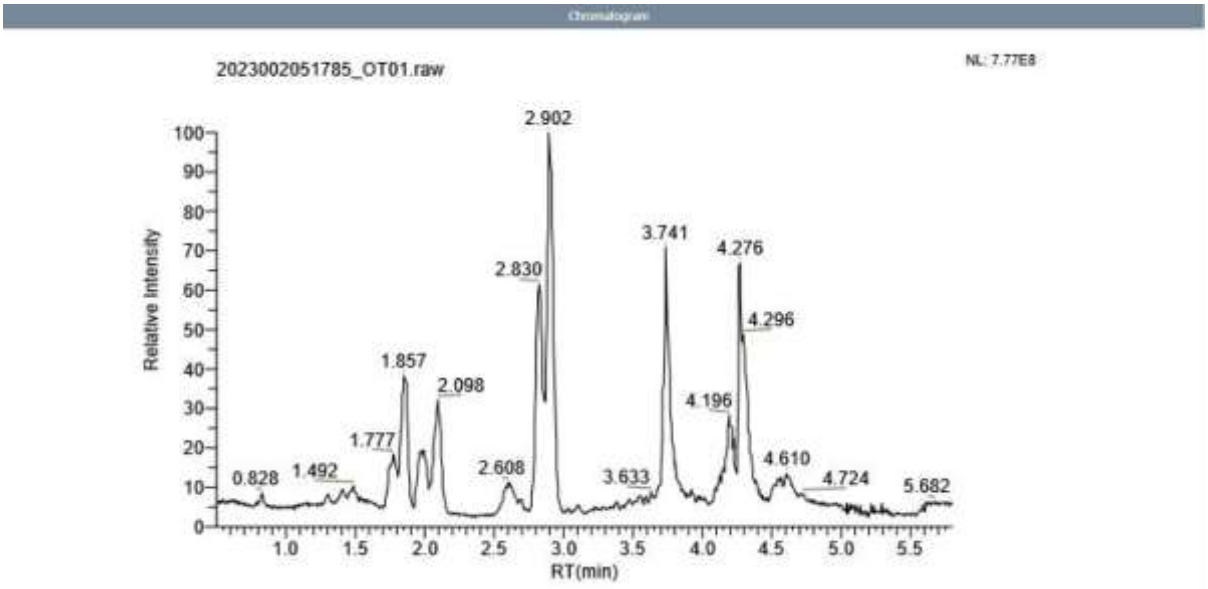


Figure SI12. LC chromatogram of PEX (**P3, T1, dT_3TP 4c, *Taq***).

RT (3.73-3.77)

(*Rows 4-5*) - DNA template T1 (calc. 8228,4713; found 8227,4833; 8238,4777; 8252,4769).

RT (4.28-4.33)

(*Rows 3, 8*) - Expected DNA product (calc. 8985,2994; found 8977,3241; 8995,3220).

Sliding Windows Xtract Masses Table

Row Number	Sequence Name	Modification	Monoisotopic Mass	Theoretical Mass (Da)	Matched Mass Error (ppm)	Average Mass	Sum Intensity	Relative Abundance	Fractional Abundance	Number of Charge States	Charge State Distribution	Number of Detected Intervals	Delta Mass	Scan Range	Start Time (min)	Stop Time (min)	Apex RT
1			1810.25264	0.00000	0.0	1811.31	309575633.74	100.00	17.21	3	1 - 3	22	0.00000	259 - 292	1.777	1.997	1.870
2			2226.06367	0.00000	0.0	2227.37	279713883.16	90.35	15.55	2	2 - 3	21	415.81134	604 - 637	4.110	4.329	4.210
3			8977.32418	0.00000	0.0	8982.35	161239020.46	52.08	8.96	8	5 - 12	18	7167.07144	622 - 650	4.229	4.415	4.282
4			8238.47773	0.00000	0.0	8243.11	158468880.27	51.19	8.81	7	5 - 11	16	6428.22510	543 - 567	3.701	3.860	3.754
5			8227.48332	0.00000	0.0	8232.11	73358342.21	23.70	4.08	7	5 - 11	16	6417.23068	539 - 563	3.674	3.833	3.734
6			2210.68662	0.00000	0.0	2211.38	70353496.36	22.73	3.91	2	2 - 3	19	399.83398	593 - 621	4.036	4.223	4.137
7			2248.64237	0.00000	0.0	2249.36	60087423.70	19.41	3.34	2	2 - 3	15	437.78973	605 - 628	4.117	4.269	4.210
8			8995.32196	0.00000	0.0	9000.36	51296253.58	16.57	2.85	8	5 - 12	34	7185.06932	625 - 676	4.249	4.590	4.302

LCMS analysis chromatogram of PEX reaction of DNA primer **P3** on the DNA template **T1** with dT_3TP **4c** and **Klenow (*exo*$^-$)** polymerase

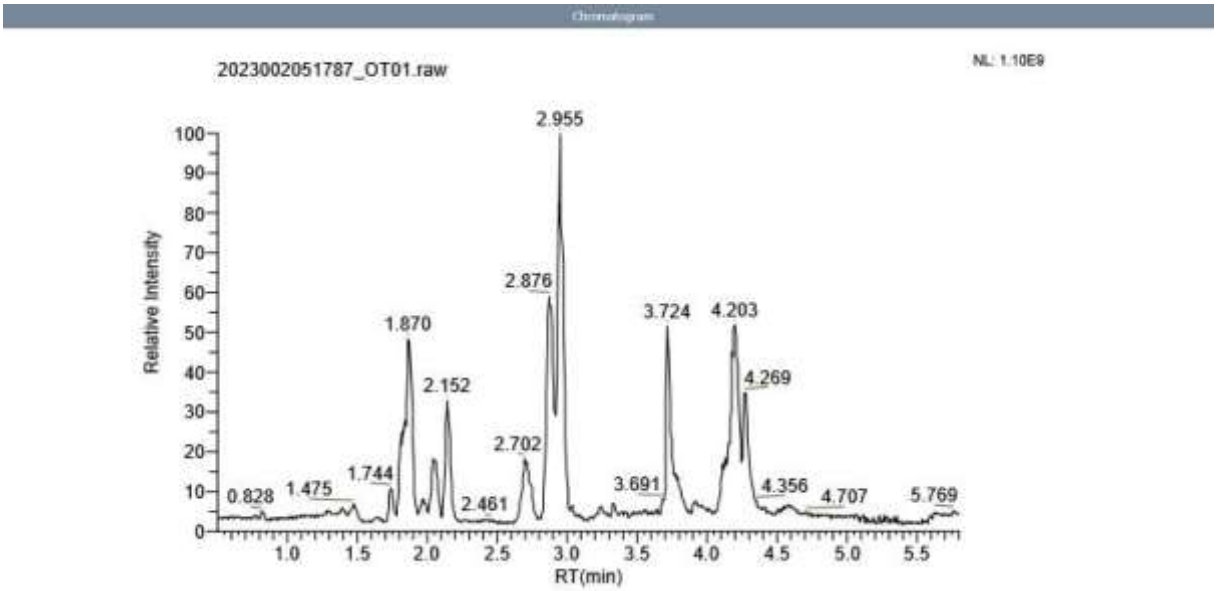


Figure SI13. LC chromatogram of PEX (**P3, T1, dT_3TP 4c, Klenow**).

RT (3.71-3.73)

(*Rows 2, 4*) - DNA template T1 (calc. 8228,4713; found 8227,4952; 8228,4968).

RT (4.20-4.31)

(*Rows 3, 8*) - Expected DNA product (calc. 8985,2994; found 8977,3019; 8995,3158).

Sliding Windows Xtract Masses Table

Row Number	Sequence Name	Modification	Monoisotopic Mass	Theoretical Mass (Da)	Matched Mass Error (ppm)	Average Mass	Sum Intensity	Relative Abundance	Fractional Abundance	Number of Charge States	Charge State Distribution	Number of Detected Intensities	Delta Mass	Scan Range	Start Time (min)	Stop Time (min)	Apex RT
1			1810.25160	0.00000	0.0	1811.31	710289462.74	100.00	42.63	3	1 - 3	20	0.00000	256 - 285	1.771	1.964	1.864
2			8227.49515	0.00000	0.0	8232.12	132168369.62	18.59	7.93	7	5 - 11	11	6417.24354	537 - 563	3.671	3.843	3.737
3			8977.30192	0.00000	0.0	8982.33	119623985.07	16.84	7.18	8	5 - 12	19	7167.05032	620 - 647	4.223	4.403	4.283
4			8228.49675	0.00000	0.0	8233.12	82469766.41	11.61	4.95	7	5 - 11	9	6418.24515	538 - 560	3.677	3.823	3.710
5			1826.23017	0.00000	0.0	1827.30	72287132.73	10.18	4.34	3	1 - 3	14	15.97056	276 - 299	1.903	2.058	1.977
6			3917.73386	0.00000	0.0	3919.97	33489396.28	4.71	2.01	5	2 - 6	13	2107.48225	480 - 500	3.285	3.420	3.346
7			3371.55625	0.00000	0.0	3373.52	32968640.47	4.64	1.98	4	2 - 5	14	1561.30463	436 - 457	2.988	3.129	3.048
8			8995.31580	0.00000	0.0	9000.35	32902883.31	4.63	1.97	7	5 - 11	29	7185.06420	626 - 670	4.263	4.559	4.393

LCMS analysis chromatogram of PEX reaction of DNA primer **P3** on the DNA template **T2** with **dTTP** and **Klenow (*exo*$^-$)** polymerase (positive control reaction)

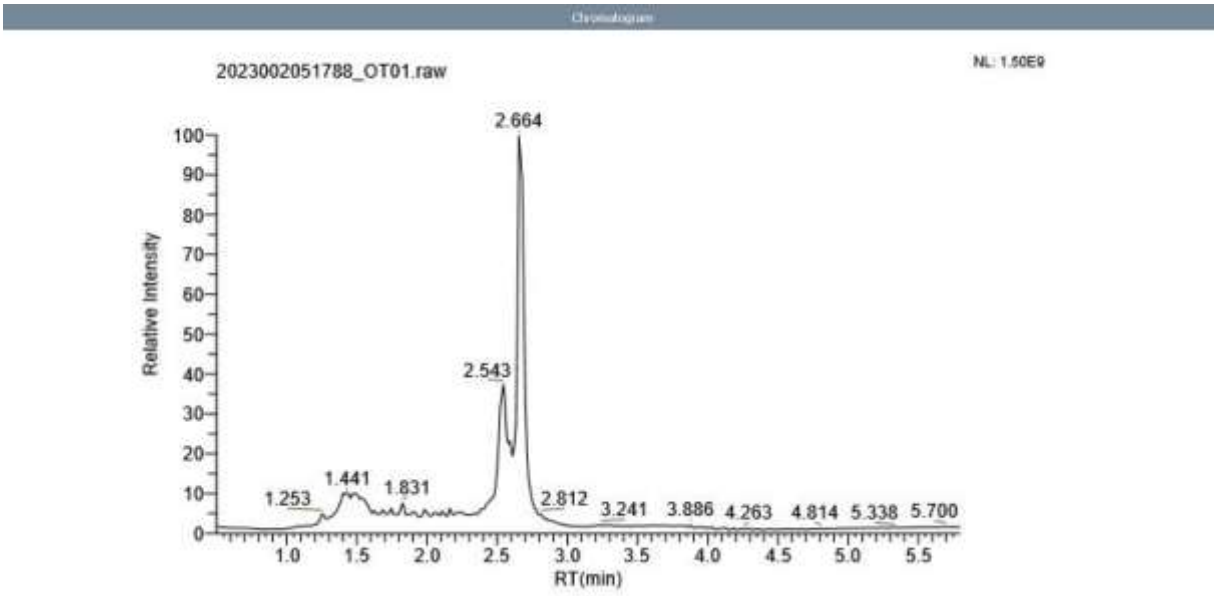


Figure SI14. LC chromatogram of PEX (**P3, T2, dTTP, Klenow**).

RT (2.53-2.57) – DNA template T2 (calc. 14406,4745)

(*Row 6*) - DNA template T2 (+GG addition) (calc. 15063,5717; found 15063,4807);

(*Rows 4*) - DNA template T2 (+GGT addition) (calc. 15367,6177; found 15367,5532).

RT (2.67-2.70) – Expected DNA product (calc. 15035,4853)

(*Row 1*) - Expected DNA product (+CCT addition) (calc. 15916,6162; 15916,5680);

(*Row 2*) - Expected DNA product (+CCTA addition) (calc. 16229,6738; found 16229,6121);

(*Row 3*) - Expected DNA product (+CTA addition) (calc. 15940,6275; found 15940,5543).

Sliding Windows Xtract Masses Table

Row Number	Sequence Name	Modification	Monoisotopic Mass	Theoretical Mass (Da)	Matched Mass Error (ppm)	Average Mass	Sum Intensity	Relative Abundance	Fractional Abundance	Number of Charge States	Charge State Distribution	Number of Detected Intervals	Delta Mass	Scan Range	Start Time (min)	Stop Time (min)	Apex RT
1			15916.56803	0.00000	0.0	15924.37	12423663.83	100.00	18.84	9	11 - 19	5	0.00000	196 - 207	2.623	2.771	2.670
2			16229.61208	0.00000	0.0	16237.57	9601512.95	77.28	14.56	9	12 - 20	5	313.04904	198 - 209	2.650	2.798	2.697
3			15940.55430	0.00000	0.0	15948.36	6443803.55	51.87	9.77	8	11 - 19	4	23.98627	196 - 205	2.623	2.744	2.670
4			15367.55316	0.00000	0.0	15375.09	4642428.21	37.37	7.04	9	11 - 19	3	-549.01487	188 - 195	2.516	2.610	2.563
5			8082.33739	0.00000	0.0	8086.30	3671417.98	29.55	5.57	6	5 - 10	3	-7804.33064	156 - 163	2.086	2.180	2.133
6			15063.48066	0.00000	0.0	15070.87	3546261.51	28.54	5.38	8	11 - 18	3	-853.08738	186 - 193	2.489	2.583	2.536
7			8395.39696	0.00000	0.0	8399.51	3292448.25	26.50	4.99	6	5 - 10	4	-7521.17107	160 - 169	2.140	2.261	2.187
8			15925.54592	0.00000	0.0	15933.35	2706678.84	21.79	4.11	5	14 - 18	3	8.97786	196 - 205	2.623	2.744	2.697
9			8724.43871	0.00000	0.0	8728.71	2234311.02	17.98	3.39	6	5 - 10	4	-7192.12932	162 - 171	2.167	2.287	2.214
10			9053.49026	0.00000	0.0	9057.92	2188133.11	17.61	3.32	6	6 - 11	4	-6863.07777	166 - 175	2.220	2.341	2.267

LCMS analysis chromatogram of PEX reaction of DNA primer **P3** on the DNA template **T2** with dT_3TP **4c** and ***Taq*** polymerase

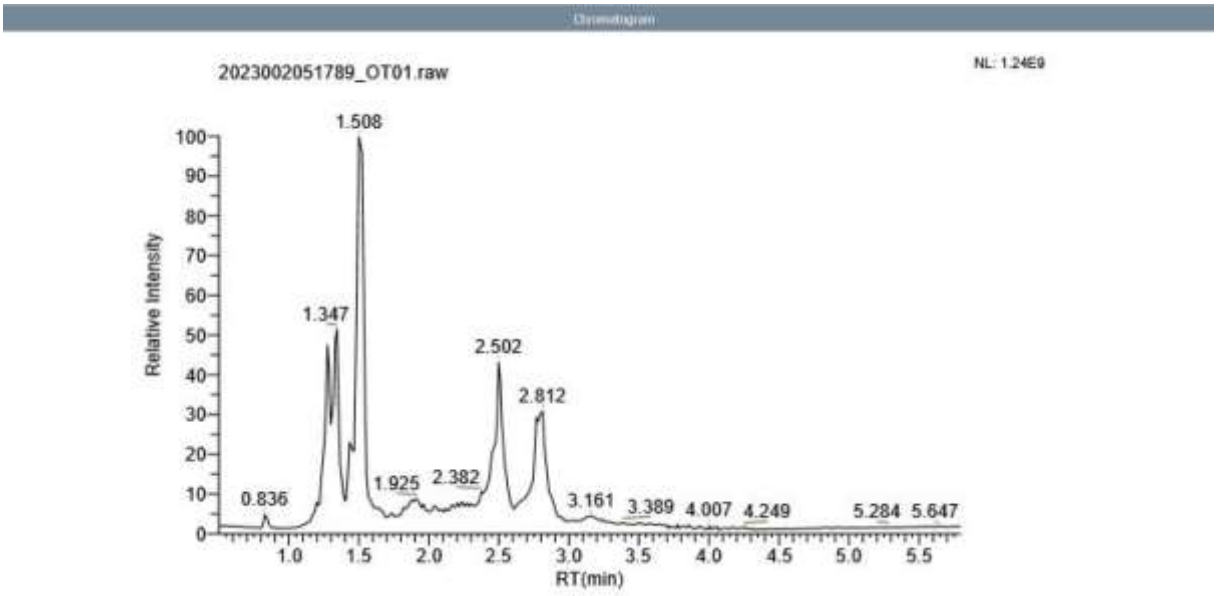


Figure SI15. LC chromatogram of PEX (**P3, T2, dT_3TP 4c, *Taq***).

RT (2.48-2.51) – DNA template T2 (calc. 14406,4745)

(*Row 5*) - DNA template T2 (+G addition) (calc. 14735,5270; found 14775,4430);

(*Row 1*) - DNA template T2 (+GGA addition) (calc. 15377,6371; found 15385,5663).

RT (2.77-2.83)

(*Row 3*) - Expected DNA product (calc. 15155,4853; found 15145,1574);

(*Row 2*) - Expected DNA product (+T*T*T addition) (calc. 16099,3431; found 16084,3163).

Sliding Windows Xcalibur Masses Table

Row Number	Sequence Name	Modification	Measured Mass	Theoretical Mass (Da)	Matched Mass Error (ppm)	Average Mass	Sum Intensity	Relative Abundance	Fractional Abundance	Number of Charge States	Charge State Distribution	Number of Detected Intervals	Delta Mass	Scan Range	Start Time (min)	Stop Time (min)	Apex RT
1			15385.56632	0.00000	0.0	15393.11	13845123.00	100.00	26.75	10	10 - 19	4	0.00000	184 - 193	2.462	2.583	2.509
2			16084.31628	0.00000	0.0	16092.20	9349394.04	67.53	18.07	10	10 - 19	11	698.74996	206 - 239	2.758	3.201	2.832
3			15145.15743	0.00000	0.0	15152.59	7703909.26	55.64	14.89	9	10 - 18	4	-240.40889	204 - 213	2.731	2.852	2.778
4			8882.33118	0.00000	0.0	8886.29	3013177.63	21.76	5.82	6	5 - 10	3	-7383.23515	160 - 167	2.140	2.234	2.187
5			14775.44302	0.00000	0.0	14782.69	2153404.28	15.55	4.16	8	10 - 17	3	-610.12330	182 - 189	2.435	2.529	2.482
6			9053.48484	0.00000	0.0	9057.92	1925518.56	13.91	3.72	6	6 - 11	3	-6332.08149	168 - 175	2.247	2.341	2.294
7			8724.44404	0.00000	0.0	8728.71	1833368.59	13.24	3.54	6	5 - 10	3	-6661.12229	166 - 173	2.220	2.314	2.267
8			8316.37251	0.00000	0.0	8320.45	1650073.69	11.92	3.19	6	5 - 10	3	-7069.19382	160 - 167	2.140	2.234	2.187
9			9366.54590	0.00000	0.0	9371.13	1519769.76	10.98	2.94	6	6 - 11	3	-6019.02043	170 - 177	2.274	2.368	2.321
10			8974.48006	0.00000	0.0	8978.88	1385710.09	10.01	2.68	6	6 - 11	3	-6411.08626	166 - 173	2.220	2.314	2.267

LCMS analysis chromatogram of PEX reaction of DNA primer **P3** on the DNA template **T2** with dT_3TP **4c** and **Klenow (*exo*$^-$)** polymerase

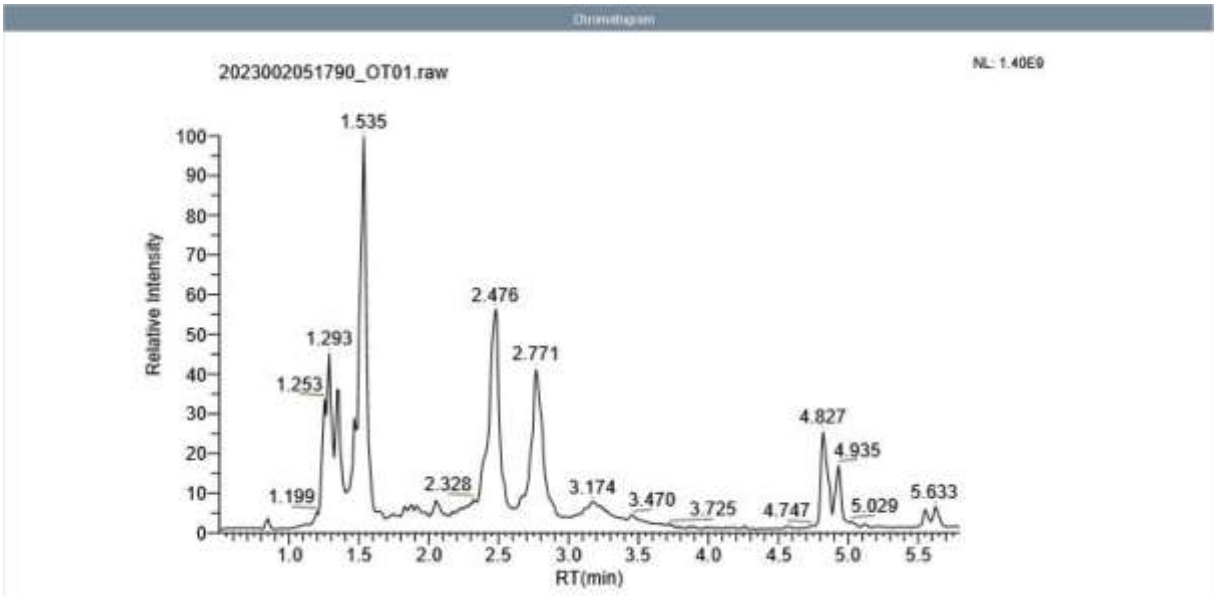


Figure SI16. LC chromatogram of PEX (**P3, T2, dT_3TP 4c, Klenow**).

RT (2.47-2.49)

(*Row 3*) - DNA template T2 (calc. 14406,4745; found 14446,3894);

(*Row 1*) - DNA template T2 (+G addition) (calc. 14735,5270; found 14759,4390).

RT (2.77-2.80)

(*Row 2*) - Expected DNA product (calc. 15155,4853; found 15145,1764);

(*Row 6*) - Expected DNA product (+T addition) (calc. 15458,2889; found 15458,2295).

Sliding Windows Xtract Masses Table

Row Number	Sequence Name	Modification	Monoisotopic Mass	Theoretical Mass (Da)	Matched Mass Error (ppm)	Average Mass	Sum Intensity	Relative Abundance	Fractional Abundance	Number of Charge States	Charge State Distribution	Number of Detected Intervals	Delta Mass	Scan Range	Start Time (min)	Stop Time (min)	Apex RT
1			14759.43898	0.00000	0.0	14766.68	11802511.00	100.00	19.76	9	9 - 18	4	0.00000	182 - 191	2.435	2.556	2.482
2			15145.17641	0.00000	0.0	15152.60	11727078.52	99.36	19.63	9	10 - 18	4	385.73743	204 - 213	2.731	2.852	2.778
3			14446.38941	0.00000	0.0	14453.48	6762609.96	57.30	11.32	9	10 - 18	3	-313.04957	182 - 189	2.435	2.529	2.482
4			14831.14357	0.00000	0.0	14838.42	3791840.63	32.13	6.35	9	10 - 18	3	71.70459	202 - 209	2.704	2.798	2.751
5			8882.35449	0.00000	0.0	8886.30	3246475.47	27.51	5.44	6	5 - 10	4	-6677.10449	160 - 169	2.140	2.261	2.187
6			15458.22952	0.00000	0.0	15465.80	3070778.79	26.02	5.14	7	12 - 18	3	698.79054	206 - 213	2.758	2.852	2.805
7			8395.39575	0.00000	0.0	8399.51	2434969.85	20.63	4.08	6	5 - 10	3	-6364.04323	164 - 171	2.194	2.288	2.241
8			9653.48779	0.00000	0.0	9657.92	2065420.41	17.50	3.46	6	6 - 11	3	-5705.95119	168 - 175	2.247	2.341	2.294
9			8724.44416	0.00000	0.0	8728.71	1883296.14	15.96	3.15	6	5 - 10	3	-6034.99482	166 - 173	2.220	2.314	2.267
10			8348.38094	0.00000	0.0	8352.47	1866387.11	15.81	3.12	6	5 - 10	3	-6411.05804	160 - 167	2.140	2.234	2.187