Snapshot of a DNA polymerase while incorporating two consecutive C5-modified nucleotides

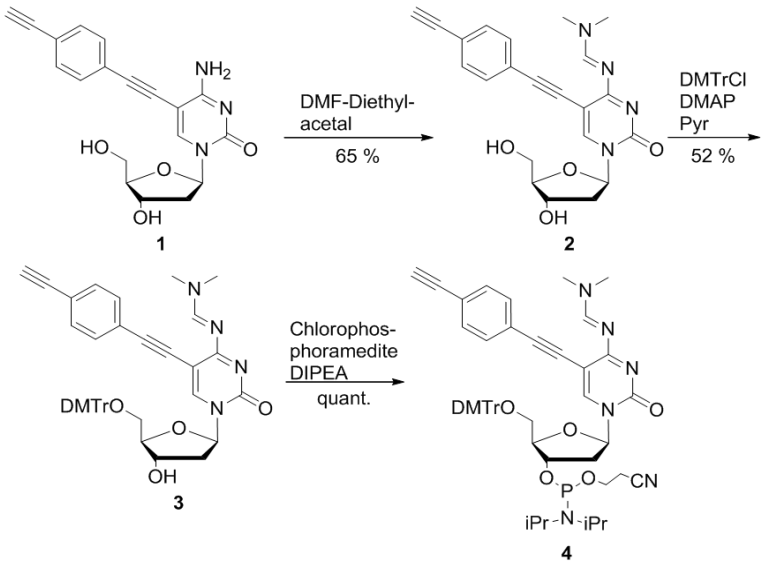
Supplementary Information

Samra Obeid, Holger Bußkamp, Wolfram Welte, Kay Diederichs & Andreas Marx

Departments of Chemistry and Biology, Konstanz Research School Chemical Biology, University of Konstanz, Universitätsstr. 10, 78457 Konstanz (Germany)

CONTENT

Nucleotide Synthesis



Scheme S1. Synthesis of dC*-phosphoramedite that is used for automated DNA synthesis.

5-(2-(4-Ethynylphenyl)ethynyl) -4-*N*-(*N'*,*N'*-dimethylformamidine)-2'deoxycytidine (2)

The nucleoside (1) was synthesized as described earlier [1]. 630 mg (1.8 mmol) of the nucleoside were dissolved in 20 ml DMF. 1.64 ml (1.32 g; 9 mmol) of *N*,*N*-dimethylformamide diethyl acetal were added. The solution was stirred at room temperature for 3 h. The solvent was removed under reduced pressure and the residue was purified by flash chromatography (12 % MeOH in DCM). 477 mg (1.36 mmol, 65%) of a slight yellow solid could be obtained.

^{1}H-NMR (400 MHz, d_4-MeOH): δ = 8.68 (s, 1H; $CHNMe_2$), 8.57 (s, 1H; H-6), 7.44 (s, 4H; ArH), 6.25 (t, $^3J(H,H)$ = 6.2 Hz, 1H; H-1'), 4.43-4.38 (dt, $^3J(H,H)$ = 6.3, 4.2 Hz, 1H; H-3'), 3.97 (q, $^3J(H,H)$ = 3.6 Hz; H-4'), 3.66 (dd, $^3J(H,H)$ = 3.1 Hz, $^2J(H,H)$ = 12.1 Hz; H-5'), 3.77 (dd, $^3J(H,H)$ = 3.6 Hz, $^2J(H,H)$ = 12.1 Hz; H-5'), 3.61 (s, 1H; C≡CH), 3.25 (s, 3H; $CHNMe_2$), 3.24 (s, 3H; $CHNMe_2$), 2.52-2.42 (m, 1H; H-2'), 2.45-2.17 (m,1H, H-2').

HRMS: m/z: calcd for $[C_{22}H_{23}N_4O_4]^+$: 407.1714; found: 407.1704

5-(2-(4-Ethynylphenyl)ethynyl) -4-*N*-(*N',N'*-dimethylformamidine)-2'deoxy-5'-*O*-dimethoxytrityl-cytidine (3)

477 mg (1.17 mmol) of the protected nucleoside (2) were dissolved in 12 ml pyridine and 15 mg (0.12 mmol) of 4-(dimethylamino)pyridine were added. The solution was stirred at room temperature over night. The reaction was quenched by the addition of 10 ml methanol. The solvent was removed under reduced pressure and was purified by flash-chromatography (2% MeOH in 1% Et3N in DCM). 428 mg (0.61 mmol, 52%) of a yellow solid could be obtained.

^{1}H-NMR (400 MHz, d_6-acetone): δ = 8.67 (s, 1H; $CHNMe_2$), 8.36 (s, 1H; H-6), 7.61-7.50 (m, 2H; DMT-ArH), 7.49-7.40 (m, 4H; ArH), 7.40-7.27 (m, 4H; DMT-ArH), 7.23-7.14 (m, 1H; DMT-ArH), 7.11-7.03 (m, 2H; DMT-ArH), 6.91-6.79 (m, 4H; DMT-ArH), 6.29 (t, ^{3}J(H,H) = 6.2 Hz, 1H; H-1'), 4.62-4.53 (m, 1H; H-3'), 4.22-4.12 (m, 1H; H-4'), 3.74 (s, 1H; C≡CH), 3.72-3.70 (m, 6H; OMe), 3.40 (dd, ^{2}J(H,H) = 10.8 Hz, ^{3}J(H,H) = 3.1 Hz; H-5'), 3.36 (dd, ^{2}J(H,H) = 10.8 Hz, ^{3}J(H,H) = 4.3 Hz; H-5'), 3.30 (s, 3H; $CHNMe_2$), 3.22 (s, 3H; $CHNMe_2$), 2.60-2.50 (m, 1H; H-2'), 2.31-2.22 (m, 1H, H-1').

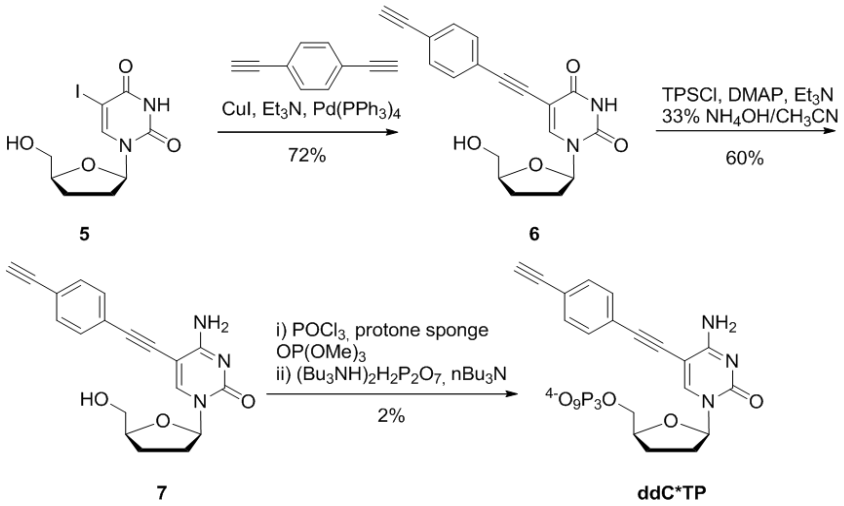
HRMS: m/z: calcd for [$C_{43}H_{40}N_4O_6Cl^-$]: 743.2631; found: 743.2639

5-(2-(4-Ethynylphenyl)ethynyl) -4-*N*-(*N',N'*-dimethylformamidine)-2'deoxy-5'-*O*-dimethoxytrityl-cytidine-3'-*O*-(c-cyanoethyl-N,N-diisopropyl)-phosphormidite (4)

100 mg (0.14 mmol) of the DMT-protected nucleoside were dissolved in 5 ml dry THF and 180 µl DIPEA. 47 µl (49 mg, 0.21 mmol) 2-cyanoethyl-*N*,*N*-diisopropylchloro-phosphoramidite were added at 0°C. After stirring for 5 minutes the ice bath was removed and the mixture was stirred for 2.5 h at room temperature. The solvent was removed at reduced pressure and the residue was purified by chromatography over a short plug of silica (1.5% MeOH in 1% Et_3N in DCM). 121 mg (0.13 mmol, 95%) of a slightly yellow oil were obtained.

^{31}**P-NMR** (162 MHz, d_6-acetone): δ = 148.22, 147.95

HRMS: m/z: calcd for [$C_{52}H_{57}N_6O_7Cl^-$]: 943.3709; found: 943.3703



Scheme S2. Synthesis of ddC*TP that is used for crystallization

5-(2-(4-Ethynylphenyl)ethynyl)-2',3'-dideoxyuridine (6)

80 mg (0.24 mmol) **1**, 90 mg (0.70 mmol) 1,4-diethynylbenzene and 9 mg (0.047 mmol) CuI (copper(I) iodine) were dissolved in DMF. The solution was degassed before 27 mg (0.024 mmol) $Pd(PPh_3)_4$ (tetrakis(triphenylphosphine)palladium (0)) catalyst was added. After a second time of degassing 66 µl (0.47 mmol) previously distilled Et_3N was added. The reaction mixture was stirred at room temperature and monitored by TLC (CH_2Cl_2 and 10% MeOH). After 2 h the solvent was removed and the product was purified by flash chromatography (solvent: CH_2Cl_2 and MeOH; stepwise gradient 0-2% MeOH) to yield 5-(2-(4-ethynylphenyl)ethynyl)-2',3'-dideoxyuridine (0.17 mmol, 71% yield).

^{1}H-NMR (400 MHz, [D_4] MeOD): δ = 8.60 (s, 1H; H-6), 7.50 – 7.38 (m, 4H; ArH), 6.04 (dd, ^3J (H,H) = 2.9, 6.6 Hz, 1H; H-1'), 4.18 (ddt, ^3J (H,H) = 3.0, 6.2, 9.1 Hz, 1H; H-4'), 3.96 (dd, ^3J (H,H) = 2.8, 12.3 Hz, 1H; H-5'), 3.71 (dd, ^3J (H,H) = 3.2, 12.3 Hz, 1H; H-5'), 3.47 (d, ^3J (H,H) = 1.5 Hz, 1H; C≡CH), 2.43 (dddd, ^3J (H,H) = 6.7, 8.4, 10.0 Hz, ^2J (H,H) = 13.6 Hz, 1H; H-2'), 2.19 – 2.09 (m, 1H; H-2'), 2.07 – 1.92 ppm (m, 2H; H-3').

^{13}C-NMR (101 MHz, [D_4] MeOD): δ = 164.0, 150.8, 145.2, 132.4, 132.1, 124.4, 123.2, 99.4, 93.0, 87.9, 83.7, 83.7, 83.6, 80.1, 62.8, 59.3, 33.9, 25.0 ppm.

HRMS: m/z: calcd for [$C_{19}H_{15}N_2O_4^-$]: 335.1026; found: 335.1034.

5-(2-(4-Ethynylphenyl)ethynyl)-2',3'-dideoxycytidine (7)

For the conversion of the modified dideoxyuridine to the corresponding modified dideoxycytidine 14 mg (42 µmol) of the modified dideoxyuridine and 8.7 mg (71 µmol) DMAP (4-(dimethylamino)-pyridine) were dissolved in acetonitrile. Under ice bath cooling 9.8 µl (71 µmol) Et_3N and subsequently 20.4 mg (67 µmol) TPSCl (triisopropyl

benzenesulfonyl chloride) were added. The reaction mixture was stirred for 2 h at 0°C. Next 1 ml of a mixture of 33% NH_4OH / CH_3CN = 1 / 1 was added to the reaction mixture. After 1.5 h stirring at 0°C the reaction mixture was allowed to warm up to room temperature and stirring was continued for another 1.5 h. The reaction mixture was diluted with CH_2Cl_2 and extracted against 1M aqueous $KHSO_4$ solution. After three times extraction of the aqueous phase with CH_2Cl_2 the organic layer were combined, dried over $MgSO_4$, and concentrated. The desired cytidine derivative (23 µmol, 57% yield) was obtained following purification by flash chromatography (solvent: ethyl acetate and MeOH; stepwise gradient 0-5% MeOH).

^{1}H-NMR (600 MHz, [D_6] DMSO): δ = 8.57 (s, 1H; H-6), 7.70 (br s, 1H; NH_2), 7.62 – 7.58 (m, 2H; ArH), 7.53 – 7.47 (m, 2H; ArH), 7.03 (br.s, 1H; NH_2), 5.90 (dd, ^3J (H,H) = 2.7, 6.7 Hz, 1H; H-1'), 5.23 (t, ^3J (H,H) = 5.0 Hz, 1H; OH), 4.32 (s, 1H; C≡CH), 4.11 – 4.05 (m, 1H; H-4'), 3.79 (ddd, ^3J (H,H) = 3.0, 5.4 Hz, ^2J (H,H) = 12.0 Hz, 1H; H-5'), 3.61 – 3.55 (m, 1H; H-5'), 2.36 – 2.27 (m, 1H; H-2'), 1.98 – 1.91 (m, 1H; H-2'), 1.87 – 1.79 ppm (m, 2H; H-3').

^{13}C NMR (151 MHz, [D_6] DMSO): δ = 163.9, 163.8, 153.4, 145.6, 132.5, 131.8, 131.4, 131.3, 123.2, 121.3, 86.5, 84.3, 84.2, 83.1, 82.6, 82.2, 61.3, 33.0, 23.8 ppm.

HRMS: m/z: calcd for [$C_{19}H_{16}N_3O_3^-$]: 334.1186; found: 334.1193.

5-(2-(4-Ethynylphenyl)ethynyl)-2',3'-dideoxycytidine-5'-triphosphate (ddC*TP)

10 mg (23.8 µmol) modified dideoxycytidine and 9.6 mg (44.7 µmol) proton sponge were dissolved in previously distilled trimethylphosphate. The addition of 3.3 µl (35.7 µmol) $POCl_3$ was carried out under ice bad cooling. The reaction was monitored by reverse phase TLC in isopropanol / water/ ammonia = 3 / 1 / 1. After 2 h stirring 293 µl (143 µmol) $(Bu_3NH)_2H_2P_2O_7$ and 79 µl (297 µmol) tributylamine were added fast and simultaneously. The solution was stirred for 2 min and afterwards quenched by adding 5 ml 0.1 M TEAB buffer. The reaction mixture was extracted three times against ethyl acetate. The aqueous solution was further purified by ion exchange chromatography The remaining sample was desalted by RP-MPLC using 0.05 M TEAA buffer and acetonitrile to afford the modified dideoxycytidine-5'-triphosphate (0.4 µmol, 1.7% yield).

^{1}H-NMR (400 MHz, D_2O): δ = 8.22 (s, 1H; H-6), 7.60 (d, ^3J (H,H) = 8.3 Hz, 2H; ArH), 7.54 (d, ^3J (H,H) = 8.3 Hz, 2H; ArH), 6.08 (dd, ^3J (H,H) = 2.6, 6.8 Hz, 1H; H-1'), 4.45 – 4.35 (m, 1H; H-4'), 4.34 – 4.25 (m, 1H; H-5'), 4.23 – 4.13 (m, 1H; H-5'), 3.59 (s, 1H; C≡CH), 3.17 (q, ^3J (H,H) = 7.3 Hz, 24H; NEt_3), 2.54 – 2.42 (m, 1H; H-2'), 2.17 – 2.05 (m, 2H, H-2'; H-3'), 1.89 – 1.87 (m, 1H; H-3'), 1.25 ppm (t, ^3J (H,H) = 7.2 Hz, 36H; NEt_3).

^{31}P NMR (162 MHz, D_2O): δ = -8.66 – -9.33 (m,1P; P_γ), -10.65 – -11.08 (m, 1P; P_α), -21.82 – -22.50 ppm (m, 1P; P_β)

HRMS: m/z: calcd for [$C_{19}H_{19}N_3O_{12}P_3^-$]: 574.0176; found: 574.0176.

Oligonucleotide synthesis

The synthesis of oligonucleotides was carried out on a solid phase DNA synthesizer (Model 392, *Applied Biosystems*) at 0.2 µmol scale applying commercially available 2-cyanoethylphosphoramidites. A standard method for 2-cyanoethylphosphoramidites was used, except for the coupling of modified nucleotides and the subsequent two nucleotides, for which coupling times were extended from 1 min to 2 min. For attachment of the modified nucleotide at the 3‘ end “universal Q-support“ (*Linktech*, UK) was used as solid support. After trityl-on synthesis the oligonucleotides were cleaved from the support and protection groups were removed by treatment with conc. ammonia at 55 °C for 12 h. After removal of ammonia the residue was purified by HPLC, concentrated, and the DMTr protecting group was cleaved by treatment with 80 % acetic acid for 2 h at r.t. The acetic acid was removed under reduced pressure and the oligonucleotides were dissolved in 100 µl water and 150 µl loading dye (80% [v/v] formamide, 20 mM EDTA, 0.25 % [w/v] bromphenol blue, 0.25 % [w/v] xylene cyanol) for subsequent PAGE purification. The DNA was detected by UV absorbance and the bands were cut out. The DNA was released from the gel by shaking the gel pieces in excess of water at room temperature for 12 h. The DNA was purified by ethanol precipitation. DNA-concentration was determined by UV absorption measurement. The oligonucleotides were analysed by HR-ESI-TOF:

Primer C*: 5’- CGA AAT TGA CCA CGG CGC*-3’

calculated mass for [$C_{184}H_{223}N_{72}O_{103}P_{17}^-$]: 5615.0 Da

found: 5615.0 Da

Competition experiment- single nucleotide incorporation assay

10 µl of the *KlenTaq* reactions contained 50 nM primer (5’- CGA AAT TGA CCA CGG CGC/C*-3’), 75 nM template (5’- ATA CAT CAA AGG CGC CGT GGT CAA TTT CG-3’), 100 µM dNTP mixture of dCTP/dC*TP in buffer (20 mM Tris HCl pH=7.5, 50 mM $MgCl_2$) and 200 nM of *KlenTaq* DNA polymerase. The ratio of modified to natural nucleotide varied from 1:20 to 100:1 (1:20, 1:10, 1:2, 1:1, 2:1, 4:1, 10:1, 20:1, 50:1, 100:1). Reaction mixtures were incubated at 37°C for 5 minutes. Primers were labeled using [γ–^{32}P]-ATP according to

standard procedures. Reactions were stopped by addition of 22.5 µl stop solution (80% [v/v] formamide, 20 mM EDTA, 0.25 % [w/v] bromphenol blue, 0.25 % [w/v] xylene cyanol) and analysed by 16% denaturing PAGE. Visualization was performed by phosphoimaging. Band intensities were read out with ImageJ.

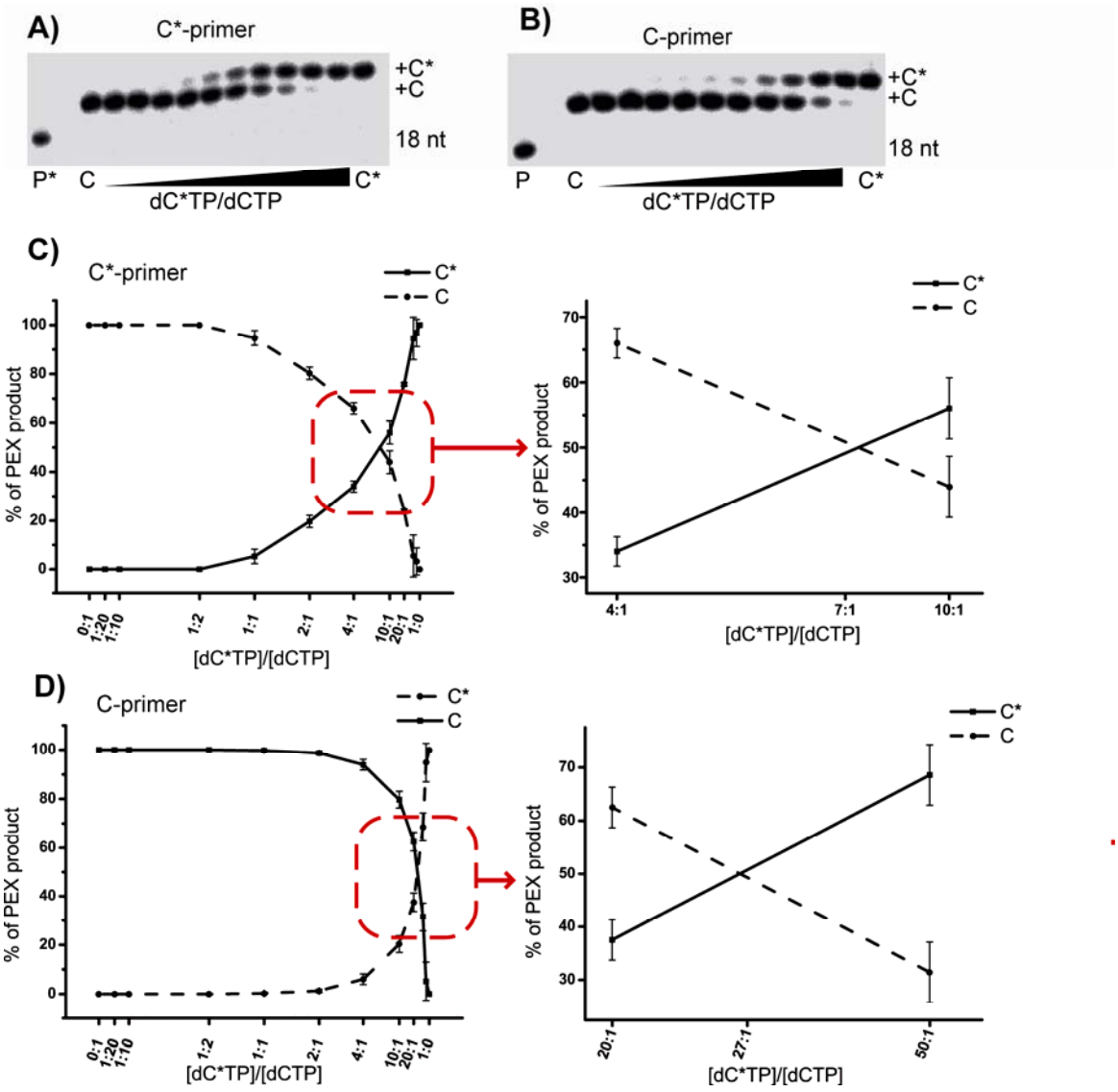
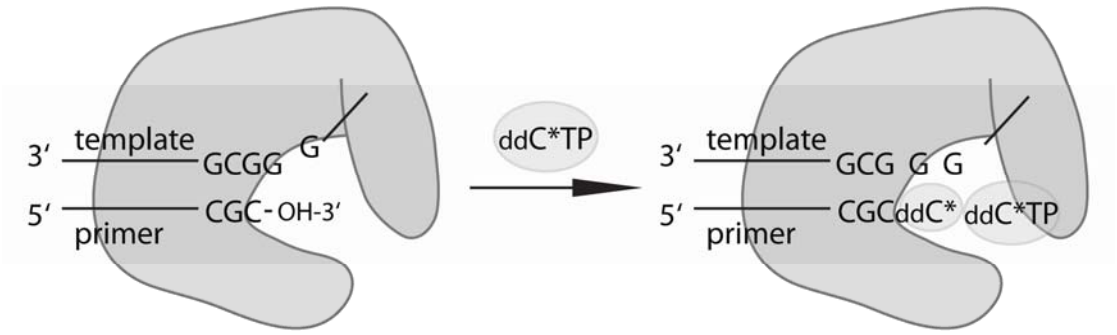


Figure S1. Competition experiments. A) A representative PAGE analysis of the competition experiment of dC*TP vs dCTP using a primer with C* at the 3' end. The ratio of dC*TP/dCTP was varied from 1/20 to 100/1 (1:20, 1:10, 1:2, 1:1, 2:1, 4:1, 10:1, 20:1, 50:1, 100:1). At a ratio of approx. 7:1 both nucleotides are incorporated with the same efficiency. B) Same as in A) beside the used primer ended with C at the 3'-end. Both nucleotides are incorporated with the same efficiency at a ratio of 27:1. C) left: The conversion in % is plotted against the ratio of dC*TP/dCTP for the C*-primer. The cross spot of both graphs is highlighted by a red dashed box. right: Zoom in to the cross spot. D) same as C)using C-primer.

Crystallization and Structure Determination



Scheme S3. Schematic depiction for the strategy to obtain DNA polymerase crystals containing two consecutively modified nucleotides.

The closed ternary complexe of *KlenTaq* was obtained by incubating *KlenTaq* in presence of DNA primer (5′-d(GAC CAC GGC GC)-3′), a DNA template (5′-d(AAA **GG**G CGC CGT GGT C)-3′), and ddC*TP. Thereby one ddC*MP is incorporated terminating the chain-reaction and a second modified ddC*TP is captured in the waiting position.

KlenTaq(ddC*): The crystallization was set up using purified *KlenTaq* (10 mg⁄ml), DNA template/primer duplex, and ddC*TP in a molar ratio of 1:3:20 and in presence of 20 mM $MgCl_2$. The reservoir solution contained 0.05 M Tris HCl (pH 8.4), 0.2 M NH_4Cl, 0.01 M $CaCl_2$, and 34% PEG 4000.

Crystal were produced by the hanging drop vapor diffusion method, equilibrating against 1 ml of the reservoir solution for 5 d at 18 °C. They were frozen in liquid nitrogen and kept at 100 K during data collection. Data of *KlenTaq*(ddC*) was measured at the beamline PXI (X06SA) at the Swiss Light Source of the Paul Scherrer Institute (PSI) in Villigen, Switzerland, at a wavelength 1.000 Å and using a PILATUS 6M detector, respectively. Data reduction was performed with the XDS package[2]. The structures were solved by difference Fourier techniques using *KlenTaq* wild-type (PDB 3KTQ) as model. Refinement was performed with PHENIX [3] and model rebuilding was done with COOT [4]. Figures were made with PyMOL [5].

Table S1: Data collection and refinement statistics

	***KlenTaq*(ddC*)**
Data collection	
Space group	P 3_1 2 1
Cell dimensions	
a = *b*, *c* (Å)	109.2, 91.1
Resolution (Å)	47.29 - 1.89 (2.00 – 1.89)
R_{meas}	11.8 (144.3)
I / σ_I	14.2 (1.20)
$CC_{1/2}$ (%) [§]	99.8 (58.0)
Completeness (%)	99.3 (95.6)
Redundancy	9.2 (6.1)
Refinement	
Resolution (Å)	47.29 - 1.90 (1.94 – 1.90)
No. Reflections	49722
R_{work} / R_{free}	17.7 (29.2) / 21.6 (31.5)
No. of residues	
Protein	539
Primer/template	12/13
Ligand/ion	ddC*TP/ Ca^{2+}/ Mg^{2+}
Water	321
B-factors (Å^2)	
Protein	38.7
Primer/template	25.8/22.9
Ligand/ion	33.1/27.0/42.4
Water	40.4
R.m.s. deviations	
Bond lengths (Å)	0.007
Bond angles (°)	1.172
Ramachandran statistics [6]	
Most favored	94.1
Additionally allowed	5.5
Generously allowed	0.2
Disallowed	0.2
PDB	4ELV

[§] $CC_{1/2}$ was calculated according to P. A. Karplus, K. Diederichs, *Science* **2012**, *336*, 1030-1033.

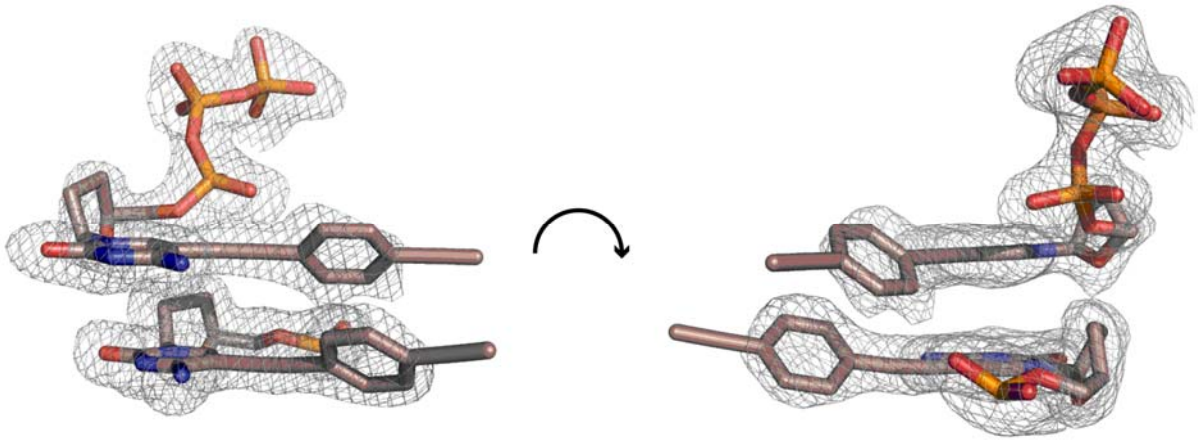


Figure S2. The final refined simulated annealing omit map mFo-DFc encountered at 3σ of the incorporated ddC*MP and the incoming ddC*TP in tow different orientations.

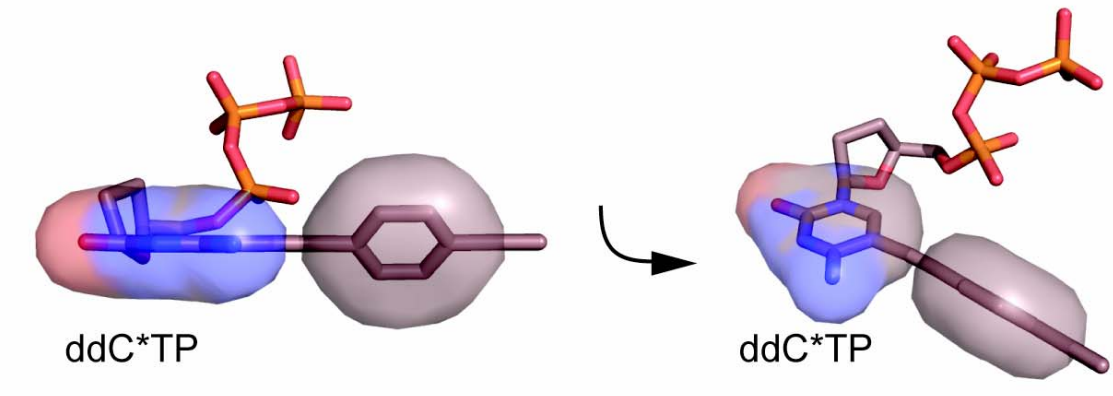


Figure S3. Two different orientations of the incoming ddC*TP of *KlenTaq*(ddC*). The aromatic ring system of the nucleobase and the phenyl ring are twisted in respect to each other.

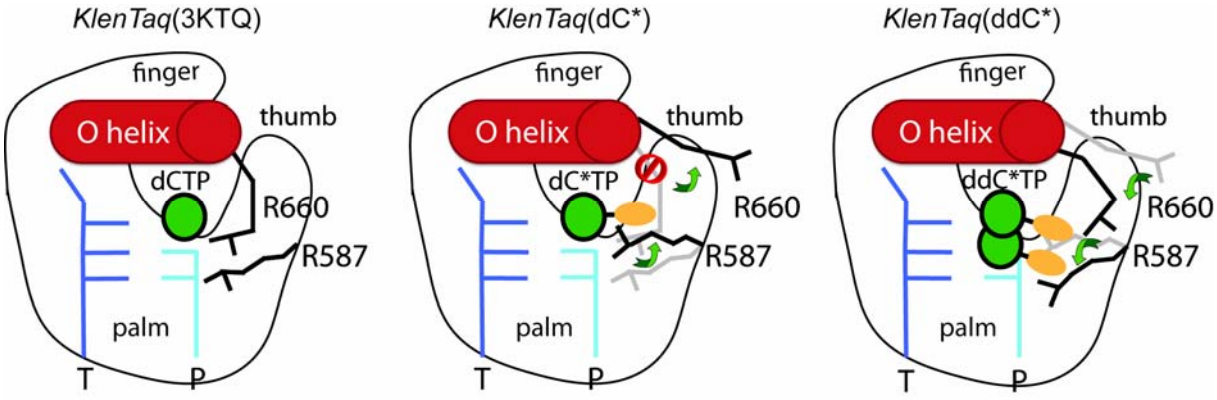


Figure S4. Schematic depiction of *KlenTaq*(3KTQ) (left), *KlenTaq*(dC*) (middle) and *KlenTaq*(ddC*) (right). Illustrated are the DNA template (T), DNA primer (P), the incoming natural or modified nucleotide triphosphates (dCTP, dC*TP, ddC*TP), the O helix (red) and the amino acids Arg660 and Arg587. The green arrows highlight the reorientation of Arg660 and Arg587. In grey are the corresponding previous position shown for Arg660 and Arg587.

References

[1] S. Obeid, H. Busskamp, W. Welte, K. Diederichs, A. Marx, *Chemical Communications* **2012**, *48*, 8320-8322.

[2] a) W. Kabsch, *Acta Crystallogr. D* **2010**, *66*, 133-144; b) W. Kabsch, *Acta Crystallogr. D* **2010**, *66*, 125-132.

[3] P. D. Adams, R. W. Grosse-Kunstleve, L. W. Hung, T. R. Ioerger, A. J. McCoy, N. W. Moriarty, R. J. Read, J. C. Sacchettini, N. K. Sauter, T. C. Terwilliger, *Acta Crystallogr. D* **2002**, *58*, 1948-1954.

[4] P. Emsley, K. Cowtan, *Acta Crystallogr. D* **2004**, *60*, 2126-2132.

[5] W. DeLano, *The PyMOL Molecular Graphics System*, DeLano Scientific, Palo Alto, CA, **2002**.

[6] R. A. Laskowski, M. W. MacArthur, D. S. Moss, J. M. Thornton, *J. Appl. Crystallogr.* **1993**, *26*, 283-291.